American Sociological Review

Volume 46 August 1981 Number 4

SCOPE AND MISSION OF AMERICAN SOCIOLOGICAL REVIEW: The *American Sociological Review* publishes work of interest to the discipline in general, new theoretical developments, results of research that advance our understanding of the most fundamental social processes, and important methodological innovations. Like other publications of the Association, emphasis is given to exceptional quality. Unlike the more specialized journals of the Association, the *American Sociological Review's* primary focus is on whatever has the most general bearing on the knowledge of society.

Concerning manuscripts, address: Editor, *American Sociological Review*, Department of Sociology, University of Illinois, Urbana, Illinois 61801. (217) 333-6515.

The *American Sociological Review* (ISSN 0003–1224) is published at 49 Sheridan Avenue, Albany, New York, bimonthly in February, April, June, August, October, and December.

Applications for permission to quote from this journal should be addressed to the Executive Office

Annual membership dues of the American Sociological Association, $20–$120 ($10 of this amount may cover a subscription to the *American Sociological Review*); Associate, $25; International Associate, $25; Student Associate, $12.

Application for membership and payment of dues should be made to the Executive Office. Subscription rate for members, $10; nonmembers, $18; institutions and libraries, $36; nonmember students, $12. Single issue, $4.

New subscriptions and renewals will be entered on a calendar year basis only.

Change of address: Six weeks advance notice to the Executive Office and old address as well as new are necessary for change of subscriber's address.

Concerning advertising, changes of address, and subscriptions, address: Executive Office, American Sociological Association, 1722 N Street, N.W., Washington, D.C. 20036.

Claims for undelivered copies must be made within the month following the regular month of publication. The publishers will supply the missing copies when losses have been sustained in transit and when the reserve stock will permit.

Second class postage paid at Washington, D.C. and additional mailing offices.

(Articles in the REVIEW are indexed in the Social Sciences Index, Psychological Abstracts, Sociological Abstracts, Ayer's Guide, University Microfilms, Abstracts for Social Workers, International Political Science Abstracts, and United States Political Science Documents.)

AMERICAN SOCIOLOGICAL REVIEW

Official Journal of the American Sociological Association

AUGUST, 1981 VOLUME 46, NUMBER 4

NEW EDITOR

Effective with the February 1982 issue, Sheldon Stryker will become the editor of the American Sociological Review. Beginning July 1, 1981, send all manuscripts to:

Sheldon Stryker, Editor
American Sociological Review
Institute of Social Research
1022 East Third Street
Bloomington, IN 47405

Notice to Contributors

A processing fee of $10 is required for each paper submitted, except for comments; such fees will be waived for student members of ASA. A check or money order, made payable to the American Sociological Association, should accompany each submission. The fee must be paid in order for the review process to begin.

The length of an *ASR* paper is generally ten (10) printed pages or thirty (30) typed pages. However, the *ASR* is happy to accept manuscripts up to fifty (50) typed pages in length for review. We also are interested in receiving manuscripts for a Research and Notes Section.

To permit anonymity in the review of manuscripts, keep identifying material out of the manuscript. Attach a cover page giving authorship, institutional affiliation and acknowledgments, and provide only the title as identification on the manuscript and abstract.

ASA Multiple Submissions Policy: Submission of a manuscript to a professional journal clearly implies commitment to publish in that journal. The competition for journal space requires a great deal of time and effort on the part of editorial readers whose only compensation for this service is the opportunity to read papers prior to publication and the gratification associated with discharge of professional obligation. For these reasons, *the American Sociological Association regards submission of a manuscript to a professional journal while that paper is under review by another journal as unacceptable.*

SUBMISSION AND PREPARATION OF MANUSCRIPTS

1. Submit four (4) copies and retain the original for your files. Copies may be Xerox, mimeograph or multilith, but not carbons.
2. Enclose a stamped self-addressed postcard for acknowledgment of manuscript receipt. *Manuscripts will not be returned.* Please do not send your original copy.
3. All copy must be typed, doublespaced (including indented material, footnotes and references) on 8½ by 11 inch white opaque paper. Lines must not exceed six (6) inches. Margins must be a minimum of one inch.
4. Include four (4) copies of an abstract of no more than 150 words.
5. Type each table on a separate page. Insert a location note at the appropriate place in the text, e.g., "Table 2 about here." Send copies, retain originals.
6. Figures must be drawn in India ink on white paper. Send copies, retain originals.
7. Clarify all symbols with notes in the margins of the manuscript. Circle these and all other explanatory notes not intended for printing.
8. Footnotes should not be placed in the body of the text. Type them (doublespaced) and attach them as a separate appendix to the text. Number them consecutively throughout the text. Footnotes are to be kept to a minimum and used only for substantive observations. Source citations are made within the text rather than in footnotes.
9. Acknowledgments, credits and grant numbers are placed on the title page with an asterisk.

REFERENCE FORMAT

A. *In the text:* All source references are to be identified at the appropriate point in the text by the last name of the author, year of publication and pagination where needed. Identify subsequent citations of the same source in the same way as the first, not using *"ibid.," "op. cit.," or "loc. cit."* Examples follow:

1. If author's name is in the text, follow it with year in parentheses. [". . . Duncan (1959) . . ."]
2. If author's name is not in the text, insert, in parentheses, the last name and year, separated by a comma. [". . . (cf. Gouldner, 1963) . . ."]
3. Pagination (without "p." or "pp.") follows year of publication after a colon. [". . . Kuhn (1970:71)."]
4. Give both last names for dual authors; for more than two use "et al." in the text. When two authors have the same last name, use identifying initials in the text. For institutional authorship, supply minimum identification from the beginning of the complete citation. [". . . (U.S. Bureau of the Census, 1963:117) . . ."]
5. Separate a series of references with semicolons and enclose them within a single pair of parentheses. [". . . (Burgess, 1968; Marwell et al., 1971; Cohen, 1962) . . ."]

B. *In the appendix:* List all source citations by author, and within author by year of publication, in an appendix titled "References." The reference appendix must be complete and include all references in the text. The use of "et al." is not acceptable in the appendix; list the names of all authors. (See A. 4. for text format.)
If there is more than one reference to the same author and year, distinguish them by the letters, a, b, etc. added to the year. [". . . Levy (1965a:331) . . ."] Give the publisher's name in as brief a form as is fully intelligible. For example, John A. Wiley and Sons should be "Wiley." If the cited material is unpublished, use "forthcoming" with name of journal or publisher; otherwise use "unpublished."
Use no underlining, italics or abbreviations.
Examples follow:

1. *Books:*

Jud, Gerald J., Edgar W. Mills, Jr. and Genevieve Walters Burch
1970 Ex-Pastors. Philadelphia: Pilgrim Press.

U.S. Bureau of the Census
1960 Characteristics of Population. Volume 1. Washington, D.C.: U.S. Government Printing Office.

Bernard, Claude
[1865] 1957 An Introduction to the Study of Experimental Medicine. Tr. Henry Copley Greene, New York: Dover.

2. *Periodicals:*

Conger, Rand
Forthcoming "The effects of positive feedback on direction and amount of verbalization in a social setting." Pacific Sociological Review.

Merton, Robert K.
1963a "The ambivalence of scientists." Bulletin of The Johns Hopkins Hospital 112:77–97.
1963b "Resistance to the systematic study of multiple discoveries in science." European Journal of Sociology 4:237–82.

3. *Collections:*

Davie, M.
1938 "The pattern of urban growth." Pp. 133–61 in G. Murdock (ed.), Studies in the Science of Society. New Haven: Yale University Press.

See recent issues for further examples.

Revised 1980 ASR

Vol. 46
No. 4-6
1981

CONTRIBUTORS TO THIS

■ MICHAEL BURRAGE (At Sixes and Sevens) is Lecturer, London School of Economics, and also teaches at the Institute of United States Studies of the University of London. He is currently studying the impact of revolution on the organization of professionals and manual workers.

■ DAVID CORRY is a lieutenant onboard Her Majesty's Canadian Ship Annapolis based in Halifax, Nova Scotia. He researched his article while studying at the London School of Economics and Political Science under the British Commonwealth Scholarship Plan. His research interests include military deviance, politics, and mass media.

■ NAN LIN (Social Resources and Strength of Ties) is Professor, Department of Sociology, State University of New York at Albany. He is studying the relationships between social resources and instrumental action and social support and illness.

■ WALTER M. ENSEL is Research Sociologist, Department of Psychiatry/Research, Albany Medical College. His current research focuses on sex, social resources, and status attainment and sex, social support, and illness.

■ JOHN C. VAUGHN is Director, Annual Surveys, Division of Research, National League for Nursing. His research interests focus on the specification and assessment of factors affecting female participation in the labor force as well as factors facilitiating upward mobility within occupation groups, particularly within nursing. He designed and annually updates the *NLN Nursing Data Book* (The League, 1979, 1980, 1981).

■ HUGH A. McROBERTS (Trends in Occupational Mobility in Canada and the United States) is Associate Professor, Department of Sociology and Anthropology, Carleton University, Ottawa. He is continuing work on status attainment and mobility, with special interest in youth unemployment.

■ KEVIN SELBEE, Carleton University, Ottawa, is involved in research on mate-selection, fertility, and labor-force participation of Israeli women; social class and income determination in Canada; and social class, mobility, and class formation in Canada.

■ J. SCOTT LONG (Organizational Context and Scientific Productivity) is Assistant Professor, Washington State University. His current work includes a cross-disciplinary analysis of department... ...ntific productivity (with ...), a study of sex-differences in ... career, and a monograph (with Robe... McGinnis), of which the current paper is a part, on processes of stratification in the scientific career.

■ ROBERT McGINNIS is Professor, Department of Sociology, Cornell University. He is continuing research with J. Scott Long into career patterns of bioscientists, especially the geographic and ecological factors that influence academic institutions in the process of recruiting bioscientists as students and employees.

■ RONALD C. KESSLER (Trends in the Relationship between Sex and Psychological Distress) is Assistant Professor, Department of Sociology, University of Michigan. He is analyzing epidemiological data focusing on the relationship between social class and mental disorders, as well as continuing research on sex and psychological distress. His recent work includes (with David F. Greenberg) *Linear Panel Analysis: Quantitative Models of Change* (Academic) and, forthcoming (with J. Ronald Milavsky, Horst Stipp and William Rubens), *Television and Aggression: Results of a Panel Study*.

■ JAMES A. McRAE, JR. is a psotdoctoral scholar with the Family and Sex Roles Program, Institute for Social Research, University of Michigan. He is currently investigating the effects of women's labor-force participation on their own and their husband's well-being.

■ ANDREW CHERLIN (Trends in United States Men's and Women's Sex-Role Attitudes) is Assistant Professor, Department of Social Relations, Johns Hopkins University. He is studying patterns of family life in the United States; his *Marriage, Divorce, Remarriage* is forthcoming from Harvard University Press.

■ PAMELA BARNHOUSE WALTERS is a Ph.D. candidate, Department of Social Relations, Johns Hopkins University. She is working on a longitudinal analysis of economic causes and consequences of educational expansion in the United States and a cross-national study of city systems and patterns of urbanization.

■ MICHAEL J. HINDELANG (Variations in Sex-Race-Age-Specific Incidence Rates of Offending) is Professor, School of Criminal Justice, State University of New York, Albany. He is conducting studies of victims of

crime and the reliability, validity, and correlates of self-reported delinquency. His most recent book (with T. Hirschi and J. Weis) is *Measuring Delinquency* (Sage).

■ JOSEPH GALASKIEWICZ (A Dynamic Study of Change in a Regional Corporate Network) is Associate Professor, Department of Sociology, University of Minnesota, Minneapolis. He is studying how nonprofit organizations secure gifts and grants from institutional donors and the effects of these interorganizational transactions on the nonprofit sector as a whole. His most recent book is *Exchange Networks and Community Politics* (Sage).

■ STANLEY WASSERMAN is Assistant Professor, Department of Applied Statistics of the School of Statistics, University of Minnesota, St. Paul. His current interests include categorical data analysis, mathematical sociology, applied stochastic processes, and joint research with Stephen Fienberg on statistical analysis of multiple sociometric relations.

ERRATA

Please see pp. 499–503 for a correction to Lizotte and Bordua, "Firearms Ownership for Sport and Protection", *ASR*, April 1980.

In the February 1981 issue of *ASR* the following errors appeared. Please remove the corrected lines printed below and insert them over the original.

Hannan and Carroll, "Dynamics of Formal Political Structure:" on p. 25, the first equation should read

$$r_{jk}(t) = \lim_{\Delta t \downarrow 0} p_{jk}(t, t + \Delta t)/\Delta t$$

Also on p. 25, the second equation should read

$$r_j(t) = \sum_{\substack{k \\ k \neq j}} r_{jk}(t).$$

Baldassare, "Effects of Household Density on Subgroups:" on p. 112, right-hand column, the corrected lines should read

(1) Household social position is divided into low status—respondents living with either the younger generation (e.g., parents living with their adult children), the older generation (e.g., nonadult and young adult children living with their parents), or respondents living with nonrelatives; and high status—which includes all other respondents not living alone.

Gove and Hughes, "Reply to Marcus and Seeman:" on p. 127, left-hand column, lines 19–22 should read

college students feared success. Although her study did include males, it was based on a very small sample and had numerous methodological problems. A number of attempts at

AT SIXES AND SEVENS: OCCUPATIONAL STATUS IN THE CITY OF LONDON FROM THE FOURTEENTH TO THE SEVENTEENTH CENTURY*

MICHAEL C. BURRAGE
The London School of Economics and Political Science
DAVID CORRY
Royal Canadian Navy

The orders of precedence issued by the City of London between the fourteenth and seventeenth centuries provide a measure of occupational status that has several advantages over modern prestige scales. A selection of the forty orders located shows dramatic changes in the status of occupations over this period, as do the records and chronicles of individual companies. Correlation coefficients suggest that both antiquity and wealth are important determinants of status, but the histories of the city companies suggest that organizational or political factors are also important. The investigation is not simply of historical curiosity; it suggests an alternative strategy for the study of occupational prestige in contemporary societies.

Once we had a quarrel on a question of precedence, Sir
And that we grew and angry too, if you will give me credence, Sirs
The Skinners were our rivals and we sought to take the lead of them;
To oust our foes at civic shows and humble all the breed of them!
This was not the toast then honoured at our dinners,
For the Lord mayor who schooled us had not yet said his say
Skinners, Merchant Taylors—Merchant Taylors Skinners.

from the "Song of the Merchant Taylors"
Margot Balfour (1927)

The archives of the city of London and the records of individual city companies provide a hitherto unused source of data for the historical analysis of occupational prestige. In this paper we draw on this material, in particular on the orders of precedence issued by the city authorities between 1328 and 1604, to analyze changes in the status position of occupations. We first briefly introduce the city of London guilds, summarize relevant aspects of their development, and then describe the circumstances surrounding the promulgation of the orders of precedence, and, on the basis of forty orders which we have collected, plot some of the changes between 1328 and 1604 in the status position of particular occupations. We next consider a few instances of status mobility in an attempt to understand the dynamics of the status system of the city over this period, and then, adopting an alternative strategy, examine the relationship between companies' positions and their wealth and antiquity. In the conclusion, we consider the relevance of this investigation to the study of occupational prestige in contemporary society.

GUILDS IN THE CITY OF LONDON

From the twelfth to the fourteenth centuries, economic life in the city of London, as in many other European cities, was organized around the craft guild. The mid–fourteenth century, in particular, marked a rapid period of growth for the guilds. Where in 1328 only 25 "mysteries" were authorized to elect their own officers, by the end of the reign of Edward III (1377) they had been joined by 35 more. Craft guilds united those who worked in a single trade or mystery, and their main functions were to control entry to the trade, to defend the occupational jurisdiction of the guild against

* Direct all correspondence to Michael C. Burrage and David Corry, London School of Economics and Political Science, University of London, Houghton Street, London, England WC2A 2AE.

trespassing by neighboring occupations or "illegal" workers, to regulate prices, materials, and the quality of products, to settle disputes between masters, journeymen, and apprentices, and generally to organize the corporate religious and leisure life of their members. There were of course many other kinds of guild, both religious and secular.

The origins of them all are obscure, though it is certain that they predate the Norman Conquest. There is documentary evidence of a "knighten guild" in London in the eighth century; this seems to have been a religious and social fraternity of landed nobility. There are records from the ninth century of the Easterlings or steel-yard merchants, a trading guild affiliated to the Hanseatic League. In many cities, though not, it seems, in London, there were also guild merchants, associations of leading merchants which endeavoured to corner and control the entire trade of their cities for their own benefit (See Herbert, 1837:3–27; Brentano, 1870:1–49). However, in the twelfth through fourteenth centuries only one other kind of guild joined the craft guilds in the city of London: the merchant guild. These, as the name suggests, consisted of merchants who did not work at the trade to which they belonged, and may not even have served an apprenticeship in it, but acted as middlemen between the producers and the customers, and whose interest, therefore, was less the manufacturing than the buying and selling of goods.

There was, and is, no public or easily recognizable sign by which to distinguish craft and merchant guilds from one another. Both were chartered in the city simply as "companies." It would, in fact, probably be more accurate to refer not to two kinds of guilds at all, but rather to two kinds of interest, the handicraft or manufacturing on the one hand and the mercantile on the other, which were present to varying degrees in all the city companies. In the earliest days, as far as we can tell, virtually all guilds were mainly composed of craftsmen who actually worked at their trade. Selling was merely an ancillary task for every master or journeyman. Over time, however, the mercantile interest asserted its control over the manufacturing. This mercantile penetration and takeover of the guilds was a piecemeal and gradual process, affecting different companies at different times. It began before the fourteenth century and continued after the seventeenth. Some companies, such as the Fishmongers', Grocers', Vintners', Drapers', and others among the twelve great companies, seem to have been dominated by merchants almost from their formation. Other companies, such as the Barbers', Bakers', Clockmakers', Chandlers', Curriers', Masons', Founders', Tiler-Bricklayers', Blacksmiths', Fruiterers', Apothecaries', and many more, seem to have remained predominantly handicraft companies long after the end of the seventeenth century.

The division of interest and of function between merchants and craftsmen means that the city companies cannot always be treated as synonymous with the occupations whose name they bear. All city companies, however, seem to have retained a close connection with the trade from which they had originated, and it seems fair to say that the numerical majority of the members of all the city companies were indeed working at the trade of their company, at least until the mid-seventeenth century.

Our evidence for this conclusion lies in the protest movements of the yeomanry of the city companies which began shortly before the Civil War, and continued during the interregnum (cf. Unwin, 1963:48). The word yeomen originally referred to the young men of a company who had completed their apprenticeship but were not yet masters. By the mid-seventeenth century, however, it described a subordinate section of the company, a few of whom might become independent masters and entitled to wear the livery of the company, but most of whom could expect to stay as journeymen for their entire lives and were altogether excluded from the government of their company. One recurrent theme in the yeomen's protests was that their companies had been taken over by "foreigners" who knew nothing of their trade (see James, 1930:ch. 5). In 1633 in a petition to the king, the yeomen of the Girdlers' Company, for example, alleged

that "of late divers merchants, silkmen, and other trades being come into the company, and bearing the chiefe offices thereof had put downe the yeomanry and appropriated to themselves the sole government of the Company, and being men of the other trades had neglected the suppression and the reforming of abuses" (Barker, 1957:56). A second example is provided by the yeomen of the Clothworkers', who in 1641 complained in a petition to Parliament that the furnishing of cloth is a "mistery" these "foreigners understand not . . . because they were never bred to it, nor instructed in it . . . . they are not sensible of governing our Mistery for they cannot be touched with our necessities . . . not being engaged in our well or our woe" (Girtin, 1958:116 ff.). Specifically, the yeomen of the Clothworkers' Company wanted a reform of the oligarchical governmental institutions of their company and demanded that the election of the master and officers be thrown open to all members of the company who were "out of their time," i.e., had completed apprenticeship. There were similar protests in many other companies, including the Goldsmiths', Drapers', and Merchant Taylors' and at least eleven of the lesser companies: the Pewterers', the Saddlers', the Girdlers' (along with pinners within this company), the Weavers', the Clockmakers', the Carpenters', the Leathersellers', the Cutlers', the Founders', the Watermens', and the Woodmongers'. In three of these, the Merchant Taylors', the Clockmakers' and the Weavers', the yeomen's protests were specifically directed against "foreigners" among the company's officers, and this suggests that though there was a considerable mercantile influence in the running of these companies, the mass of the companies' membership were still following the trade which gave the company its name. The yeomen's protest would otherwise be meaningless. For their part, the officers of these companies obviously accepted the legitimacy of the yeomen's argument, for they replied, as the Clothworkers' Court of Assistants did, that they, or most of them, were not foreigners at all, but had learned their trade as apprentices and journeymen and were doing their utmost to uphold the rules of their trade and protect its interests (Girtin, 1958:117–8).

If, therefore, the city companies in the mid–seventeenth century were not completely synonomous with a single trade or occupation, one may reasonably conclude that they were still tied to the trade whose name they bore, and that the defense of an occupational jurisdiction was still one of their major preoccupations. Since merchant domination of city companies tended to increase over time and was particularly marked among the twelve great livery companies, we can be more confident that the boundaries of city companies coincide with those of a single occupational group when we consider earlier dates and companies other than the twelve great ones.[1]

In setting up their own governing institutions, trades in the city of London typically went through four distinct stages. At first their organization was entirely informal; members of the trade had rules of apprenticeship, regulated the supply of materials, set prices, and elected officers to enforce them. During this period their activities were unauthorized and members of these "adulterine" guilds risked prosecution and fines. Eighteen of them were fined in 1179 (Unwin, 1963:47). There is evidence though, that these fines were imposed on an annual basis like a tax or fee, which suggests that the guilds were tolerated by the authorities. During this first stage of their development, the guilds came closest to the fraternity of equals idealized by Durkheim and Ruskin, among others. In each of the later stages their internal structure became more hierarchi-

[1] Kellett (1958) discovered a livery list of 1756 which gives the proportion of the membership of thirteen companies actually following the nominal trade of their company. In nine of these companies the proportions vary between 76% and 93%, in one it is 36%, and in the three remaining 8%, 7%, and 6%. All of the companies listed were minor ones, and Kellett observes that membership of the twelve great companies was much less directly linked with practice of the company's nominal trade. These figures, however, relate to a time more than one hundred and fifty years after the close of the period we are considering. We have failed to locate any similar list compiled during our period. We suspect there is none because the issue was not then the problem it had become by 1756.

cal and their government more oligarchical.

In the second stage the trade was incorporated, usually by an ordinance of the court aldermen of the city, though a number of companies obtained charters directly from the monarch and thereafter usually had them confirmed at the beginning of every reign. More exceptionally, a few obtained charters from Parliament. Charters sometimes included valuable trading privileges, but in itself the grant of incorporation probably did not lead to fundamental changes in the running of the company. Even before incorporation, most companies elected their own masters and wardens and effectively governed their own trades. City ordinances and royal or parliamentary charters therefore only legitimized powers that many companies had previously exercised, though no doubt legal recognition also made it easier for the company to acquire property, to hold its courts, to assess quarterage from its members, to conduct searches through the city, and to fine delinquent members. No doubt, too, the charters helped companies defend themselves against other companies trespassing on their occupational jurisdiction.

The third stage was the grant by the city authorities of a livery—that is, permission for officers and members of the guild to wear a distinctive dress. The grant of a livery was a recognition of the status of a guild, but it was by no means purely honorific. Only livery companies participated in the political life of the city, since the common council, which may be roughly compared with a lower and rather weak legislative chamber, was elected solely from their members. Many civic offices were also restricted to the members of livery companies.

The fourth stage was the grant of a coat of arms by the king's official, the Lord Garter of Arms. This was a final status accolade and merely provided adornment for companies which had already achieved a secure position. In addition, there were many other honors which were bestowed upon a privileged few companies. Edward III, in a deliberate attempt to raise the guilds in public estimation, himself became "a brother" of the Linen-Armourers', later re-christened the Merchant Tailors'. Thereafter, many of his successors, other members of the royal family, dignitaries from the church or the nobility became honorary members. By 1607 the Merchant Tailors' had had as honorary members "seven Kings, one Queen, two Duchesses, five countesses, and two baronesses, seventeen princes and dukes, one archbishop, one and thirty earls, and a hundred other lords and gentlemen" (Herbert, 1837:29). In this practice we can detect the beginning of the gradual transformation of the city companies from economic bodies concerned with making or selling goods into purely social and honorific associations concerned with administering estates, dining, pageantry, and good works.

These four stages summarize the typical status-progression of city companies, as well as their constitutional development. Not all companies, of course, followed this typical path. The Parish Clerks', for instance, obtained ordinances from the city but were granted neither a livery nor a royal charter. Some of the livery companies never received or, apparently, sought a royal charter. Others had a royal charter but no livery. Our summary gives, so to speak, the main road. Many companies deviated from it, but those companies that reached the fourth stage were usually to be found at the top of the status hierarchy, while those merely incorporated by city ordinance were at the bottom.

One further point must be mentioned to complete our introduction to the status order of London. Twelve livery companies managed, sometime between the late fourteenth and mid-fifteenth century, to distinguish themselves, both in terms of power and status, from all other companies whether liveried or not. Exactly how this came about is obscure, and there seem to be no parallels to this segregation in other medieval cities. In any event, after the death of Richard II (1399), thirteen companies began to separate themselves from the rest and were thenceforward, in Herbert's words, "always dignified by some epithet implying *superiority,* wherever they are mentioned: as 'the substantial companies; the principal

crafts; the chief mysteries; the wealthier and superior companies; the most worshipfull felowshippers,' and other similar titles" (Herbert, 1837:37–8). Twelve of these companies, later known as the twelve great companies, maintained this exclusive position as a sort of separate status order and have done so to the present day. The lord mayor and the court of aldermen, the effective governing body of the city, were elected solely from the livery of these twelve companies. Their distinctive status was recognised at all public functions where they invariably took precedence. At coronation ceremonies, for example, only the wardens of these twelve companies attended the lord mayor; they alone (with the exception of the Linen Armourers') could enroll the sovereign as a member, and generally only they could receive foreign princes and ambassadors (Herbert, 1837:37–8, fn.).

THE ORDERS OF PRECEDENCE

So far we have considered only the typical status "careers" of the companies. We turn now to the status relationships between companies, which is the main subject of our paper. From 1328 on, the status position of every city company was fixed and published in the city's orders of precedence. These orders determined the position of each company in the civic activities of the city as well as their contribution to its political and economic life. The orders are therefore an excellent source of data for studying status stratification in London.

In the period with which we are concerned, there were an enormous variety of civic activities. Many of these involved the monarch and his family: in particular, coronations and royal marriages or funerals, but including royal "visitations" to the city. The city also celebrated military victories and had its own functions and holidays, chief among them being the installation of a new lord mayor. There were also a large number of religious feast days. For all of these events, the participation of every company was determined by its position in the order of precedence. On royal visits, for example, each comapny was assigned a "station" on the main streets, on which the members were expected to erect flags, banners, and a decorated stand for themselves and their guests. From this, incidentally, the phrase "above or below one's station" seems to originate. Each company was similarly assigned an exact position in river pageants, processions, and feasts, and was required to contribute, according to its position, horsemen, guards, entertainment, provisions, or money.

Apart from these civic pageants and celebrations, there were many other occasions when the city required the participation of the companies. At various times it was necessary to mount a guard at the city gates, and the city regularly required financial contributions for military expenditure, for public works in the city, such as building the Guildhall and Royal Exchange, for widening streets, or for digging sewers. Other assessments were for charities, chosen by the monarch or by the city. Much more significant were the assessments to provide funds for the monarch's military campaigns and daily expenditure, a practice which Parliament adopted during the Civil War. Thus, in searching through the city archives one finds a considerable number of lists giving the order of precedence of the city companies. By collecting and assembling these, we may present a historical profile of the status order of the city of London from the mid–fourteenth to the beginning of the seventeenth century. After this date, the status order was fixed, apart from reshuffling following the disbandment of companies. Why the status order of the city froze in this way is still an open question, though it seems likely that it was a consequence of the progressive decline in the companies' economic functions.

Before describing the results of our investigation it must be said that these orders of precedence vary in validity, i.e., in the extent one can say that they actually are orders of precedence rather than lists of the companies compiled in a more or less random manner for administrative purposes. Some commentators, for example, have questioned whether the nine lists reproduced by Herbert in 1837 really record the status position of the companies, and the question arises a fortiori with re-

spect to our much larger collection (Herbert, 1837:101–3). In our view the validity of the orders of precedence falls into three categories. In the first, there can be no doubt whatever that the list is intended to set out an order of precedence, simply because in introducing the list, the city authorities explicitly refer to the fact. The list of January 31, 1516 has pride of place in this respect since it was specifically and solely intended to settle many disputes over precedence which had, it seems, disturbed city life in previous years. The preamble to this list is as follows:

> Here after ensuyth thorder and direccion taken at this court by the mayr and aldremen above seyd of and for all the craftes and misteres ensuying ffor their goyinges as well in all processions as all other goynges Standynges and Rydynges for the busyweissys and causes of this Citie, The seyd order and direccion to be from hensforth fermely observed and kept eny other Rule order or direccion heretofore made to the contrary notwithstanding provided alweys that the ffelysship whereof the maire ys for the yere accordying to the olde custumme shall have the preeminence in goyng afore all other ffelysshippes in all places durying the tyme of mayralte etc.

There follow the names of forty-eight companies headed by the Mercers (*Repertory of the Court of Aldermen* 3 f. 66b, 31 Jan. 7 Henry VIII). This 1516 list, along with that of 1488, has face validity.

In a second category we would put a number of lists which are not solely concerned with matters of precedence but which, due to their context, are almost certain to be in order of precedence. On February 18, 1463, for example, the court issued a list which it introduced as follows:

> This day it is agreed that at the next coming of the King from the north parts of the City, the mayor and aldermen shall meet him on horseback in scarlet and that the commoners shall also meet the King on horseback dressed in blue gowns with black bonnets and bripipes, and that from all the underwritten companies certain persons shall ride just as is determined below, that is to say:

There follows a list of thirty-nine companies with the numbers of men each will provide (City Record: Journal 7, fo. 21b, trans. from Latin by P. Edwardes). A second example is the list issued on June 24, 1509, when the city promulgated an order of crafts for the coronation of Henry VIII, which is introduced with the following words:

> Hereafter apperith the ordo(u)r of crafts howe they shall stond when the Kyng and the Queyn shalle passe by towards their coronaction the chef crafte to begyn on the south side in the chepe a little from the olde chaunge ende and so furth toward the towne. And ev(er)y of them to make Rayles at their p(ro)p(e)r costs for their standings after the novmbre (folowyng) [City Record, Journal 10, fo. 370b.].

In our view the validity of such lists as orders of precedence is reasonably certain.

There remains a third category which includes lists which do not refer to status either directly or indirectly. These are not concerned with public ceremonies, where precedence is manifestly a matter of some consequence, but with some rather mundane aspect of city administration. One issued on August 24, 1485 begins, "These craftes as folowyng be assigned to watche nyghtly from VII of the clok at nyght unto VI of the Clok in the mornyng & that they shall mete at Yeldhall and there vaite uppon ii aldremen &c." (City Record, Journal 9, fo. 84.) Other examples are the list of 1586 for the provision of gunpowder and that of 1598 for a loan to the Queen "for defence of realm and suppression of Rebels in Ireland." (City Record, Journal 24, fo. 324b.) The validity of such lists as orders of precedence may reasonably be questioned.

For several reasons we are inclined to think, however, that lists of this third category are indeed orders of precedence. First, it is inherently improbable that the officers of a city, all of whose activities and communal affairs were imbued and directed by questions of precedence, would, when compiling lists for a loan, for an assessment of money or grain, or for a watch at the city gates, so forget themselves that they would list the companies in any old order or casually demote companies whom they knew to be very jealous of their position in the order of precedence. The clerks of the city were, we know, rather slapdash and inconsistent in

their spelling, but it seems to us implausible that they were also casual and inconsistent about the order in which they listed their companies. In this context the lists compiled after 1604 are of particular significance, since after 1604 we know exactly what the order of precedence of the city companies was and do not have to rely on the list itself for guidance. We can therefore check whether lists issued for administrative purposes do or do not follow the order of precedence. An example is the ordinance issued by the lord mayor and aldermen in 1665 which instructed the companies to lay up a yearly supply of coal "for the use of the poor in times of scarcity and to defeat the combinations of coal dealers" (Herbert, 1837:130–1). This list clearly is for purely administrative purposes and has nothing explicitly to do with status positions, yet the list of companies in this ordinance follows exactly the order of precedence. Our belief is that the city clerks followed the same practice before 1604 as after 1604, routinely listing companies in their correct status order. Moreover, if the administrative lists issued before 1604 are not in order of precedence, we are at a loss to explain what order they are in. They are not alphabetical, nor are they by size of assessment, since companies high on the list frequently have smaller assessments than those below them. Nor are they random, because there is, despite changes in the position of individual companies, a marked continuity over time in the order in which companies are listed. Since those orders correlate highly not only with each other but also with civic honors and responsibilities, we are strengthened in our belief that customary practice in the city was to list companies in their order of precedence, irrespective of the purpose for which the list was issued.

Following this line of reasoning, we have included lists of all three categories in Table 1, though since there may be some remaining doubt about those issued for administrative or fiscal purposes, we have also provided an indication of the validity of each list. We have done this in two ways, first by rating the validity on the basis of our reading of the introduction to the list and the events which surrounded it according to the three categories above. A score of 1 is given for those lists which are issued purely as orders of precedence. Those which surrounded an event, such as processions and river pageants, of importance to the crafts and in which precedence clearly mattered, are given a score of 2; the lists arising from purely administrative matters are assigned a score of 3. Our second method of measuring the validity of the lists is to correlate the rank order of each list with those immediately preceding and immediately following (in the case of 1328, only the latter, of course, is possible) on the assumption that drastic changes in the order would not be made without the question of precedence being debated either in the city or in the company concerned. Thus, if the reliability coefficient were under .90, it would give one prima facie grounds for questioning the validity of the list. Three lists have reliability coefficients below .90: 1328, 1422 and 1463. Those of 1422 and 1428 both refer to important ceremonial occasions. There is, therefore, reason to believe that, despite their low reliability coefficients, both are indeed orders of precedence. The list of 1328 has low validity by both of our measures and therefore remains suspect. It certainly appears that precedence was at this time less structured and organized than it later became, and the city clerks may not have strictly adhered to it in compiling this list of "those elected and sworn in divers misteries of London, for the government and instruction of the same."

We have located forty lists in all, and since these lists contain up to 89 companies, we obviously cannot reproduce them all, even in an abbreviated form.[2]

---

[2] We cannot say for sure that these forty lists include every single order of precedence or list giving an order of precedence issued by the city authorities during our period. To be certain of that would require a complete search of the entire city archives, a formidable undertaking. It is, however, doubtful whether the inclusion of further lists would significantly alter the general picture, and in any case we can be pretty sure that we have included the most important lists, for the simple reason that such lists are frequently referred to in the annals of the city companies and are thus repeatedly drawn to our attention.

Table 1. Selected Orders of Precedence of City Companies 1328–1604, with order in 1837

Order number	1	2	3	4	5	6	7	8	9	10	11	12	13
Year	1328	1422	1463	1470	1488	1509	1516	1532	1545	1586	1596	1604	1837
Validity	3	2	2	3	1	2	1	2	3	3	3	3	1
Reliability	.50	.66	.86	1.00	.89	.98	.94	.93	.99	.97	1.00	1.00	—
Companies in original list	25	31	39	64	65	46	48	60	28	51	55	55	89
Companies in table	14	22	25	25	25	24	25	25	22	25	25	25	25
FISHMONGERS	1	6	4	4	4	5	4	4	4	4	7	6	4
GOLDSMITHS	2	9	5	5	6	6	5	5	5	5	11	7	5
DRAPERS	3	3	2	3	3	4	3	3	3	1	4	5	3
GROCERS	4	2	3	1	2	3	2	2	2	3	3	4	2
IRONMONGERS	5	8	9	10	11	8	10	9	10	10	10	10	10
SADDLERS	6	10	22	15	15	18	24	20	19	16	24	24	25
MERCERS	7	1	1	2	1	2	1	1	1	2	2	3	1
GIRDLERS	8	20	12	21	21	22	22	24	21	18	22	22	23
VINTNERS	9	5	6	8	10	11	11	8	11	11	9	11	11
HABERDASHERS	10	21	11	11	8	9	8	10	8	8	5	1	8
BREWERS	—	11	14	24	4	13	14	21	14	20	14	14	14
BUTCHERS	11	12	15	22	22	17	23	22	—	22	23	23	24
LEATHERSELLERS	—	—	13	25	25	16	15	13	15	24	15	15	15
WAX CHANDLERS	—	13	20	18	18	—	19	17	—	25	19	19	20
SKINNERS	12	4	7	6	7	7	6	6	6	7	1	8	7
BAKERS	—	—	16	23	23	14	18	23	—	23	18	18	19
TALLOW CHANDLERS	—	14	19	19	19	23	20	18	16	13	20	20	21
PEWTERERS	—	15	21	13	13	20	16	14	18	21	16	16	16
DYERS	—	16	18	12	12	12	13	12	13	14	13	13	13
ARMOURERS & BRASIERS	—	17	24	17	17	24	21	16	22	19	21	21	22
SHEARMEN[a]	—	18	23	20	20	15	12	19	12	12	12	12	12
SALTERS	—	19	10	7	9	10	9	11	9	9	8	9	9
CUTLERS	13	—	17	14	14	19	17	15	17	15	17	17	18
MERCHANT TAYLORS	14	7	8	9	5	1	7	7	7	6	6	2	6
BARBERS[b]	—	22	25	16	16	21	25	25	20	17	25	25	17

SOURCES: See Appendix.
[a] In 1528 the Shearmen united with the Fullers to become the Clothworkers (Unwin, 1904:44ff.).
[b] Until 1745 the Barbers were united with the Surgeons, but they are usually listed in the orders of precedence simply as the Barbers.

We have therefore selected for Table 1 twelve orders of precedence issued between 1328 and 1604. Our criteria of selection were, first, completeness of the list and, second, a relatively even spread of lists over the period we are considering. In reality, a good number of lists are clustered round particular dates. Eight lists, for example, were issued between 1483 and 1488, three in 1509 and four between 1586 and 1588. A random selection might therefore provide an even more incomplete picture than that presented in the table. We have also included the order of precedence compiled for the Royal Commission on Municipal Corporations in 1837. This is essentially the same as the order of precedence at the present time. Our table covers the twenty-five companies which headed the 1837 order of precedence and therefore includes all of the twelve great livery companies; the rows record the changes in each company's position in the order of precedence relative to the other twenty-four companies. They do not record their true ordinal position in the city's orders of precedence. If we had attempted to give the true ordinal positions of all the companies on each list, we would have given a misleading impression of changes in the status position of companies, since many of the changes would have been a function of variations in the total number of companies listed. By selecting an almost constant set and describing only status movements of companies relative to other companies in the set, we can be more confident that we are plotting real status movements. The set is not wholly constant because in the lists of 1328 and three other years. Some of our chosen twenty-five companies are not mentioned.

The third and fourth rows at the top of Table 1 give the two measures of validity mentioned above: our validity rating and

the reliability coefficient, while the fifth row records the total number of companies included in the city's original list. The source of each list is given in the appendix.

The main conclusion to be drawn from the table is that there were significant movements in the status position of the city companies between the fourteenth and seventeenth centuries, especially among the lesser companies in the lower half. The table clearly shows the superior status of the twelve great companies, there being no movement into or out of the top twelve positions from the mid–fifteenth century apart from the Shearmen/Clothworkers. In the present context the lower half of the table is the more important, for the twelve great companies were, as we have already observed, dominated by merchants for much, if not all, of the period; and we therefore cannot say that their position in the order of precedence gives the status ranking of the occupation whose name they bear. For the lesser companies, we can, however, be more confident of this conclusion.

STATUS MOBILITY

Our next task is to breathe some life into the figures in our table by describing the circumstances surrounding the status mobility of some of the companies. One sort of movement can be dealt with quite briefly, i.e., movement in and out of the number one position. It was, as the preamble to the 1516 order quoted above observes, an "olde custumme" of the city that the company to which the lord mayor belonged was automatically placed first during his year of office. The other types of movement revealed in the table are, however, more difficult to understand. While the records and histories of the city companies often refer to the status ambitions, insecurities, and conflicts of their subject they rarely give the reasons for their promotion or demotion, and there is almost no evidence about the deliberations of the court of aldermen when they arbitrated in questions of precedence. Their decisions were simply announced without explanation, and we are left to wonder about the "dyverse conderacions" that they often referred to in announcing their decision and about the backstairs lobbying and politicking that probably accompanied their deliberations. There are, however, some references to status disputes which had entered the public domain, which, together with the available evidence of events surrounding some of the more conspicuous instances of status mobility, provide insight into the dynamics of the orders of precedence.

One may note, to begin with, that the evidence dispels any suspicion that the city orders of precedence were a mere formality for public events or simply a matter of administrative convenience. Disputes over precedence provoked passionate feelings and at times escalated to violence and even armed conflict. In 1226 a disagreement between goldsmiths and tailors, apparently over matters of precedence, left many killed and wounded in the city streets. The combatants were separated only by the sheriff at the head of the *posse comitatis*, and the ringleaders were later executed (Noorthouck, 1773:56). These two companies again fought pitched battles in 1267–1268 and were joined by some clothworkers and cordwainers. Once again there were many dead and wounded and twelve of the ringleaders were executed. The Goldsmiths' seems to have been a particularly belligerent company. In 1325 they were again involved in violent conflict, this time with the Saddlers'. There was also an armed conflict between fishmongers and skinners in 1340, which similarly provoked much bloodshed and ended with the execution of the leaders. (For details of these and other disputes, see Unwin, 1963:59–60, 81, and Herbert, 1837:25, 100). Herbert unearthed a manuscript from 1375 which records, "This yere at Awrestlynge [wrestling] John Northwold, mercer was sclayn at the black heth, where thorough aroos a gret discension and debate among the craftes of London." The author does not further explain this remark, and Herbert merely adds, "It seems, that the like sort of contention for superiority which had existed in the reign of Henry III had now separated them [the city companies] into direct opposite parties" (Herbert, 1937:35).

Armed conflicts between members of the city companies are rarely found in the period we are considering, though there are numerous references to "affrays" and "scuffles" in the histories of the companies where matters of precedence were involved. In some cases companies sought to assert their status claims, if not violently, then at least physically, by occupying the station in civic events to which they thought they were entitled, rather than that to which they had been assigned by the city authorities. In 1427, for example, the Butchers' sought to usurp the station of the Goldsmiths'. As the historian of the Goldsmiths' Company describes it: "In matters of precedence feelings always ran high and . . . the company fought and won a bitter dispute with the butchers who had sought, on the newly-sworn mayor's return from Westminster, to place themselves along Goldsmith's Row in Cheapside, evicting the company from its excellent and traditional vantage point" (Reddaway, 1975:107).

The Barbers' were involved in protracted status disputes in the first half of the sixteenth century. In 1516 they were adjudged twenty-eighth company, where previously, by their own account, they had been placed seventeenth. They protested, but nothing was done to redress their grievances until February 1532, when they somehow prevailed on the court of aldermen to give them back their former position, and the lord mayor directed one of his sergeants to wait on the Pewterers', the eighteenth company and "shewe theym in that the seyd company of Barbours Surgeons be Restored ageyn to their olde Rowme" (City Record, Rep. 8, Leaf 272b, 6 Feb. 24 Henry VIII (1532)). This decision seems to have provoked considerable protest from other companies, for three months later, in May 1532, the Barbers' were assigned eighteenth place by the city authorities. They appear, however, to have had some difficulty in holding on to this position, and Young suggests that rival companies may have ignored the city's orders and forcibly ousted the Barbers (Young, 1890:239–41).

One year later the Barbers' appealed again to the city authorities to be restored to their old place of seventeenth company. Their petition was accepted, and on this occasion we can catch a rare glimpse of the reasoning behind a decision of the court of aldermen. Their proclamation dealing with the matter declared that the

> mayr and Aldermen consydering not onlye the sayd Request to be good and Reasonable but also the good qualyties and humanytie whiche the sayd companye have and shewe from tyme to tyme yn alman [er] Tasks contrybucyons and other charge[e]s borne and levyed of of [*sic*] amongs the sev[er]al occupacons of this cytie wheryn they be found always Ryght tractable redye and conformable and also for-asmoche as yt appereth by tholde Records w[i]t[h]yu this Cytie that they have used to be ye the said XVIIIth Rowme as our theyr behalf ys afore Alledged. Therefore and for dyv[er]s other causes and considerac[i]ons theym specyally movyng . . . A full Co[u]rte of Aldremen . . . with good delyb[er] acyon and advysement fully agreed and gr[a]unted that from hensforth at all times to come forev[er] more the master wardeyns and companye of the sayde mysterye for the tyme beyng shalbe accepted taken and Admytted the XVIIth companye . . . [City Record, Letter Book 0, Leaf 213–213b, 4 Feb. 24 Henry VIII (1533)]

This decision apparently provoked further protest from the companies humbled by the elevation of the Barber-Surgeons'; and as a result, a little over a year later the company was pushed back to twenty-eighth place. To stop the company again trying to assert their claim by physically occupying a higher position, the court of aldermen ordered that the barber-surgeons "shall no more goo yn p'cessyons, standynge, Rydynge, goynge, and other assembles from hensfurth, tyll it be otherwyse ordered by thys corte" (City Record, Rep. 9, Leaf 79, 24 Oct. 26 Henry (1534)).

After this rebuff the Barber-Surgeons' did not abandon their claim, though it is not clear from their annals exactly how they prosecuted it. After a further year they once again occupied seventeenth position, and were confirmed in this position four times in 1535 and twice in 1536. Despite further attempts to dislodge them in 1604 and 1606 and the shuffling of positions when the Stockfishmongers' united

with the Fishmongers' and when the dyers were elevated from eighteenth to thirteenth place, the barbers have retained this position to the present day (Young, 1890:241).

As we observed earlier, it is unfortunate, though perhaps understandable, that we know much more about the status quarrels that led to public confrontation than about those that were settled by lobbying, influence, or negotiation in private. Clearly, in the case of the Barber-Surgeons', the crucial activities were mainly behind the scenes. It seems likely that there was also intense politicking in what is perhaps the most celebrated squabble over precedence in the history of the city of London, that between the Merchant Taylors' and Skinners'.[3] The city first decided to intervene in this long-running dispute in 1483–4, when the two companies quarrelled as to which of them should take precedence in the lord mayor's procession to Westminster. The lord mayor and aldermen resolved the dispute by decreeing that they should, in alternate years, take sixth and seventh positions; and it is from this decision that the phrase "at sixes and sevens" originates. In an attempt to promote harmony between the two companies, it was also ordered that they should dine together at each other's hall once every year. This attempt at reconciliation failed, and the squabble went on for a further thirty-eight years, with the Skinners' claiming that the 1484 ruling applied only to civic processions and "general goings and assemblies," but that in all other ceremonies and events they had permanent precedence. In 1521 the court of aldermen ruled against this claim, and decided that precedence was to alternate between the two companies in all public business. So it has remained until the present day.

One reason why the order of precedence could never remain fixed for very long was that economic development led to the emergence of new occupations and therefore of new companies which the city authorities then had to place in the order of precedence. This process is evident in our table. In the 1328 list it may be seen there were only twenty-five companies and that their number tended to increase over time (though not all city companies are included in every list), so that by the end of the fifteenth there were more than sixty. New trades were seldom, of course, entirely new. They were usually formed by the members of an existing company who had begun to specialize in one aspect of their trade. Their specialization slowly became recognized, by them at least, as a distinct trade or mystery. They then tried to distinguish and protect themselves from their erstwhile colleagues and sought independent incorporation and official recognition as a separate company which required a position in the order of precedence.

Apothecaries provide an interesting illustration of this process. By the early seventeenth century they had long formed a distinct group within the Grocers' Company, specializing in the preparation and sale of medicines and often providing the only medical attention and advice available to the poor. Some members of the company, of course, still sold medicines along with groceries, and apothecaries had not succeeded in establishing themselves as a separate mystery. After years of petitioning, however, the full-time apothecaries within the Grocers' Company finally obtained a royal charter of incorporation in 1617, and the question therefore arose of their position in the order of precedence. This was not settled for thirteen years. By an act of common council in 1630, they were placed "next before the Painter-Stainers'" in twenty-eighth position. The company immediately protested against this decision, arguing that "His Majesty never intended that they should be cast to the twenty-eighth company . . . considering their arte which is not Mechanick but Liberall." They also pointed out that a committee of aldermen which had earlier considered the question had recommended that they be placed after the Leather-Sellers', i.e., in sixteenth position (Wall and Cameron, 1963:23–5, 28–9).

We can only speculate as to the reasons for the common council's decision. The

[3] For details of this dispute see City Record folios 195b–196, 10 Apr. 1 Richard III (1484); Repertory 5 fo. 166b; Clode, 1875:243–5; and Herbert, 1837:100.

Apothecaries' Company, being new in 1630, was of course still small and poor and could not indulge in the conspicuous consumption of wealthy, well-established companies. City companies were intensely proud of their halls and furnishings, their silver plate, their banquets and entertainments, their barges, and so forth, and vied with one another in displaying these status symbols. The fledgling Apothecaries' Company had little to show in this respect, and it may be that the city authorities gave them a low position for this reason. One suspects, however, that the wealthy merchants who dominated the Grocers' Company, and were well represented in the court of aldermen and common council, may have had something to do with the decision, since it is known that the Grocers' Company, second in the city order of precedence at the time, resented and resisted the apothecaries' struggle for independence.

HYPOTHESES

Our remaining task is to explain, or try to explain, the status position of the companies in the orders of precedence. In 1837 the Royal Commission inquiring into the municipal corporations suggested that "precedence seems to have been regulated by the period of their incorporation, which in general corresponds to their importance" (Second Report of the Commissions Appointed to Inquire into the Municipal Corporations in England and Wales, London, 1837). Reddaway feels that "prestige, numbers and civic importance was the yardstick," while Herbert (1870:101) decided that there was "no fixed principle" to determine precedence. Blackham (1931:82) refused to speculate about causal factors, stating that it was "neither antiquity nor wealth" that determined order of precedence, "nor was it numerical strength, but it has now been accepted for so long that it seems idle to speculate on the subject."

We do not agree with Blackham's conclusion that it is idle to speculate on the subject, though we cannot claim that our speculations have led us to any definite conclusion. We have however, tested two hypotheses: first, that order of precedence is determined by antiquity, and second, that it is determined by wealth. As a measure of antiquity we have taken the date of earliest incorporation of the various livery companies and have correlated this with the companies' position in the orders of precedence, using both Pearson's r and Spearman's rho, though, since our data are ordinal, the latter is the more appropriate measure. The results for eight of the orders are presented in Table 2. The selected lists include those where the correlations between antiquity and position are the highest and the lowest and also provide a relatively even spacing over the period, so that if there were any change in the relationship between antiquity and position in the order of precedence it would be evident. We have also given the measures of reliability and a validity rating of each list, as defined in the second section of this paper.

The data presented in Table 2 indicate that antiquity, as measured by date of incorporation, is related to order of precedence but that a significant variation in the order of precedence is left unexplained. Moreover, the correlations we have found must be interpreted warily, for antiquity,

Table 2. Date of First Charters of Incorporation as Related to Orders of Precedence, London 1422–1837'

Date of order of precedence (N)	Validity	Reliability	Pearson's r	Spearman's rho
1422 (27)	2	.66	.651	.702
1465 (42)	2	.91	.736	.780
1485 (47)	2	.98	.552	.677
1509 (42)	2	.95	.547	.692
1516 (42)	1	.94	.533	.688
1540 (52)	3	.97	.589	.705
1598 (50)	3	1.00	.583	.711
1837 (76)	1	1.00	.730	.819

' NOTE: For operational definitions of reliability and validity, see second section, "The Orders of Precedence."

or date of incorporation, is not a wholly independent variable. Many of the early incorporations were able to secure valuable trading privileges in return for loans and support to the monarch, and it may be, therefore, that it was the trading privileges they had obtained rather than their antiquity that helped these companies to maintain their status position. From the seventeenth century on, precedence was of course determined solely by antiquity. The Company of Airline Pilots, to take an instance from more recent times, became a livery company in 1956 and was given eighty-first place in the order of precedence, the last vacant position at the time. However, in the period which we are considering, it is clear that antiquity cannot be the whole story.

The second hypothesis is that order of precedence is determined by the wealth of the companies. Examination of the orders of precedence themselves makes it clear that status was not synonymous with wealth. When making their assessments, the city authorities took into account the relative wealth of the companies and their capacity to pay, and these do not coincide with their status position. As noted earlier, low-status companies are frequently assessed for greater amounts of money, wheat, coal, etc. than high-status companies. The Stationers' are an example of a low-status company of great wealth. They, Herbert (1870:38) observes, "besides their growing wealth and extensive concerns rank higher as a rich, commercial and working company" than the twelve great companies. There is also evidence that some of the wealthier but lesser companies would at times gladly have paid more than the amount they were assessed by the city authorities. The twelve great companies reserved for themselves the privilege, as they saw it, of contributing to the cost of the repair of the city wall irrespective of the willingness or capacity of lesser companies to contribute. All these facts confirm that status was not determined solely by wealth. If that had been the case, there can be little doubt that the status mobility of the companies would have been far greater than it was. This being said, it is also clear from the orders of precedence that there was some relationship between wealth and status; the higher status companies were generally wealthier than the lower. Furthermore, as we noted earlier, it would appear that when the economic functions of the companies declined and disappeared, status movements also came to an end and the order of precedence was frozen.

To obtain a measure of the relative wealth of different companies, we have taken the sums each company was assessed when the city was raising money for a "loan" to the monarch or for charitable purposes. Table 3 gives the correlation, in both Pearson's r and Spearman's rho, between the ranking of the company and the amount assessed for the ten lists containing such information. It also gives the reliability measure of each list, though not the validity score. Since these lists were compiled for fiscal purposes, they all fall into our third and lowest category of validity.

The results presented in Table 3 show a high correlation between wealth as we have measured it and order of precedence. Wealth therefore seems to be a better predictor of status than antiquity. This correlation, however, is also ambiguous, since wealth, like antiquity, is not an independent variable. A company's economic position was influenced by the type of charter obtained, and this in turn was influenced by political circumstances at the time of incorporation and the bargaining skills of the company in wringing privileges and concessions from the king or Parliament. The wealth of city com-

Table 3. Individual Company Contributions to Various Loans and Charities, as Related to Orders of Precedence for Given Years; London 1506–1665

Date of order of precedence (N)	Reliability	Pearson's r	Spearman's rho
1506 (36)	.89	.678	.776
1545 (27)	.99	.901	.927
1573 (28)	.96	.865	.840
1586 (49)	.97	.770	.828
1591 (53)	.92	.742	.805
1596 (53)	1.00	.767	.794
1598 (53)	1.00	.750	.827
1602 (53)	1.00	.787	.785
1604 (54)	1.00	.760	.798
1665 (55)	.80	.728	.718

Table 4. Date of First Charters of Incorporation and Company Contributions Related to Orders of Precedence 1506–1665 (by First Order Partial Spearman's e rho)

Date of order of precedence (N)[a]	Company contributions by precedence controlling for date of charter	Date of charter by precedence controlling for company contribution
1506 (34)	.773	−.109
1545 (24)	.919	.515
1573 (27)	.717	.543
1586 (44)	.770	.089
1596 (47)	.730	.302
1598 (47)	.752	.565
1602 (47)	.684	.170
1604 (48)	.714	.205
1665 (49)	.603	.247

[a] Numbers differ from preceding tables because satisfactory record on both variables is required.

panies is therefore to some unknown degree the result of successful political action.

Since neither antiquity nor wealth alone can fully explain the orders of precedence, we attempted to measure the combined effect of these factors by means of partial Spearman's rho. This seems the most suitable measure of the relationships between three rank orders, though for a number of reasons we think the results should be treated with considerable caution.[4] Multivariate analyses require a precise definition of variables, and our operational indices of both antiquity and wealth are, as we have already suggested, not above suspicion. Moreover, the data on the indices themselves is incomplete and unsatisfactory. Some charters were lost in the Great Fire, and many ancient companies were incorporated by means other than a royal charter. Thus we hesitate to draw firm conclusions from the calculations which are presented in Table 4.

Bearing these reservations in mind, the results suggest that the relationship between wealth, as measured by the size of a company's contribution, and the order of precedence remains strong when controlling for the date of incorporation. By contrast, the relationship between precedence and the date of incorporation is reduced substantially in almost every case, suggesting that the correlation in the earlier zero order is partially due to its interactive effect with company wealth. In sum, wealth seems to be strongly related to order of precedence, while antiquity is only related as it interacts with wealth. However, the figures also suggest that some other factor or factors are responsible for a company's position in the order of precedence. Our mini-case studies confirm this. It would be absurd, for instance, to suppose that the shuffling of the Barber-Surgeons' between seventeenth and twenty-eighth position reflected changes in the company's economic fortunes. Nor can it have anything to do with its date of incorporation.

We are prompted therefore to suggest a third factor to help explain the order of precedence, namely political action. As we noted earlier, it seems reasonable to infer that behind the scenes in the barber surgeons' dispute there must have been a great deal of bargaining, lobbying and negotiation. It seems likely that this also occurred in many other cases. Though difficult to document, any attempt to understand changes in the status position of companies would, we think, have to take into account this political aspect of status. Status position, in other words, was not simply granted or given by the city authorities on the basis of wealth, antiquity, and other attributes, but had to be asserted, striven for, and pursued over long periods of time. Status seems to have depended, in part at least, on the organizational skills of the company, on the degree to which they mobilized support among their own members and powerful allies, as well as on their ability to mollify or outmaneuver their potential rivals.

DISCUSSION

Our investigation raises a great many questions which we have neither the time nor adequate empirical evidence to answer. We have mentioned, for instance, some of the behavioral correlates of status position: specifically, the way in which

[4] Since the theory of partial rank order correlations is less developed than that of interval data, this is a debatable proposition. Reynolds (1971) obtained some encouraging results using different partialling procedures, including Spearman's rho, for a number of ordinal measures; we have followed him here.

position in the order of precedence determined companies' participation in the political and ceremonial life of the city and their civic responsibilities. We have said nothing, however, about interpersonal status-relationships and the way relationships between individual freemen were affected by the relative standing of their companies, as when a freeman of a high-status company had dealings with a Master or Warden of a low-status company. It may be that there was another status order in the city which overlay that given in the orders of precedence, and which placed the Masters, Wardens, and members of the courts of all companies in a separate and superior category to the freemen of all companies. To identify this, however, and understand the way it interacted with company status would require a different kind of search from that which we conducted in the city records. We would only note that the yeomen protesting about the government of their companies in the seventeenth century did not make common cause against their "betters," as we might have expected them to have done if they had recognized themselves as belonging to a deprived and inferior stratum of city society. Though roughly contemporaneous, all of these protest movements seem, curiously enough, to have been quite independent of one another.

Further research would also illuminate the social distance between companies, about which we have been able to say very little. Since the twelve great companies shared certain political and ceremonial privileges, it seems reasonable to suppose there was more social contact between the members of these companies than between them and any of the minor companies. We have, however, found little evidence that would enable us to compare the social distance of the members of the adjacent minor companies with that of the members of companies some way apart in the order of precedence. Table 1 presents only a fraction and is biased, since it covers only those companies that were of the top of the order in 1604, though at earlier times many of them had been much lower in the orders of precedence. It would be interesting to conduct a similar analysis of all 55 companies in the list of 1604 to see if one could detect status "bands" within which the minor companies move, and which would be informally comparable to the distinction between the twelve great companies and the rest. In this context it would be particularly interesting to follow the fortunes of nonmanual occupations, such as the scriveners, minstrels, innholders, stationers, parish clerks, and apothecaries, all of whom are usually to be found in the lower half of the orders of precedence. In the prestige hierarchies of modern societies the distinction between nonmanual and manual occupations is often thought to be a fundamental dividing line; but in the orders of precedence of the city of London in our period it is, apparently, of little significance.

There is, finally, still a great deal to be learned about the motivations behind companies' struggles to better their position in the order of precedence. Why did the Merchant Taylors' and Skinners' quarrel for so long? Why did the Barber-Surgeons' put up such a struggle? What was the prize? Was it simply self-esteem, or was there something else at stake? There are, here and there in the city records, some indications that the charters of high-status companies provided better protection for their trade than the charters of those beneath them in the order of precedence. Moreover, it seems that the regulatory authority of the higher companies, and in particular their authority to search and fine, sometimes extended over the members of lesser companies. There are cases where members of inferior companies were brought before the courts of more senior companies. Knowing more about the extent of these practices would help us understand the motives behind the struggle for precedence, and might enable us to assess the relative weight of the political, economic, and status rewards of the struggle.

All these questions will, however, have to wait another occasion. Our aim in this paper has been simply to introduce a new source of data for the historical study of occupational status, to present the results of our initial investigations, and to consider possible explanations of the status mobility we have discovered.

Since there are few other status systems, apart from other medieval cities, where relative positions of occupations are determined with such precision, where there is a recognized arbiter of status disputes whose decisions are publicly announced and strictly enforced, it may appear that this investigation is of purely historical interest.[5] We would reject this conclusion and will make two observations in support of this view.

By far the most common method of measuring occupational prestige in contemporary societies is by reputational scales. These scales have been used to make comparisons over time and, more commonly, between societies. We also have made comparisons of occupational prestige, or something close to it, over time. However, we have presented our results, or a part of them, in the form of a table showing status movements. Comparative studies of prestige in modern societies, whether making comparisons over time or cross-culturally, rarely present their conclusions in this manner, and would perhaps find it difficult to do so. Instead they compare prestige scales by correlating them with one another and on the basis of the resulting coefficients decide how similar or dissimilar the scales are. Treiman (1977:29–41), for instance, correlated 83 scales from contemporary societies, plus one from fourteenth century Florence and another from nineteenth century Nepal. Since, with very few exceptions, the scales were all highly correlated with one another, he concluded that, the hierarchy of occupational prestige was invariant for all complex societies past and present.

Out of curiosity we decided to apply this "comparison by correlation" approach to our data and analyzed the intercorrelations between all 40 of the orders of precedence. The intercorrelations were high. Of 37 lists correlated with the order of precedence fixed by the court of aldermen in 1516, 21 had a Pearson's r of .9 or better, 7 of .8 or better, 5 were .7 or over, and the other 4 listings were of the period 1425 prior to the institutionalization of orders of precedence. Had we adopted this method, therefore, we would have concluded that the order of precedence of the city of London companies had been static for over half a millenium.

This conclusion we know to be false. Our table, limited as it is to about a quarter of the companies, tells us it is false. Abundant evidence from the annals of city companies also tells us it is false. Manifestly, therefore, the "comparison-by-correlation" method is at fault. It simply does not pick up significant changes over time. This method would not even have enabled us to distinguish between the dynamic period in the orders of precedence and the period from the late seventeenth century to the present day when they were indeed frozen. In our view this is because Pearson's r, which Treiman uses, and which we have also used for certain purposes, assumes homoschedacity, and the data violates this assumption: i.e., a few extreme scores, which are highly correlated, bias the correlation coefficient despite widespread variations among those at the center. Thus, while many companies in the city of London were constantly changing their status position, a few dominant companies, and in particular the twelve great livery companies, had a relatively stable position; and as a result, because of the disproportionate weight attached to the extremes, the correlation coefficients "prove" that the status order of the city was more or less static, but fail to detect the mobility of the mass of occupations in the middle.

This obviously raises doubts about the supposed uniformity and constancy of occupational prestige in modern societies. Anyone reading the documentary evidence about the status striving and competition between occupational groups in contemporary societies, about their sense of insecurity and their responses to it, must feel uneasy when they also read that sociologists, using a generally accepted method of measuring and comparing oc-

[5] The closest comparable status systems are those of the Florentine guilds in the fifteenth century (The Guilds of Florence, 1906:61–2) and of the castes of Nepal from 1395 (Wright, 1877:111–2), both as cited by Treiman (1977:116). There are also some points of resemblance between the city companies and status of castes in India (Dumont, 1970:134–8) and the Table of Ranks of prerevolutionary Russia (Pipes, 1977:124–5).

cupational prestige, have concluded that the status position of various occupations is, to all intents, constant. Our investigation can only increase this feeling of uneasiness.

Our second observation relates to the method of construction of contemporary prestige scales. These scales are constructed by outsiders, in the sense that they are derived from the opinion of a random sample of respondents about a selected number of occupations, in which they are not involved nor, one imagines, particularly interested. The orders of precedence, by contrast, were constructed by the actors themselves, in the sense that the companies put their case to the common council, and the members of the council, who themselves all belonged to city companies, deliberated and decided, presumably after considering the opinions and objections of other companies. This difference in the method of construction would appear to give the orders of precedence a decided advantage over reputational scales, and it leads us to suggest a possible alternative strategy for the study of occupational prestige in modern societies.

This strategy would begin by studying occupational prestige, not in entire societies, but within selected occupational families such as medicine, the law, public administration, industry, education, and the military, in a manner somewhat similar to that suggested by Hatt (1950) in his critique of prestige scales thirty years ago. He identified occupational "situses" such as the professions, agriculture, business, and the military and within them occupational "families"—e.g., within the professional situs he distinguished the families of the free professions, the pure sciences, the applied sciences, and community professionals. By contrast, the occupational families we are proposing share a common work environment and interact with one another in the performance of their work. They resemble in some respects, therefore, the status order of late medieval London.

One can distinguish, for example, a medical status order in modern societies, involving specialists, professors, hospital administrators, housemen, nurses, general practitioners, medical technicians, auxiliaries, white-collar staff, ambulance drivers, and hospital maintenance and catering workers, among others. Apothecaries, as it happens, have participated in both medieval and modern status orders. Having failed to make much headway in the orders of precedence of the city of London, the apothecaries transferred to the status order of modern medicine and devoted most of their attention to improving their status relative to other occupations involved in medical care: physicians, surgeons, managers of workhouses, poor-law administrators, and local government officers. Their modern successors, the general practitioners, are still engaged in the same enterprise, though with a different set of adversaries.

Studying status within particular occupational families would seem to have decided advantages over the analysis of status via the reputational scale administered to the population at large. Like the citizens of medieval London, the participants in particular occupational families have a pretty clear notion of both their own status position and that of all the other participants in it. Their ratings of the relative prestige of occupations are therefore likely to be more informed and precise than those of random samples of the population. Moreover, they also know the status distance that separates them from everyone else. They know the behavioral implications of their position—the way they are expected to behave in the presence of their superiors or inferiors—and are aware of the status claims and strategies of occupations with whom they are in close working contact. It therefore seems likely that investigations of status within closely related occupations would provide more accurate prestige scales and more comprehensive accounts of status in societies than studies based on the responses of the ignorant, uninvolved, and largely indifferent public at large. It also seems likely that such investigations would provide better leads to the dynamics of status stratification, because most, if not all, of the significant status conflicts in modern societies are between neighboring occupations, e.g., between different kinds of engineers, between dif-

P2020

ferent kinds of medical professional or paraprofessional workers, or between different kinds of industrial workers. Such conflicts would be highlighted in the kind of investigation we are proposing, whereas in national studies of occupational prestige these conflicts are hidden by amalgamating such occupations into single occupational categories or by juxtaposing them in prestige scales with occupations with which they have little or no contact and can therefore rarely, if ever, be in dispute. Studies of status within particular occupational families would, in short, pinpoint status rivalries and conflicts, which national scales of prestige can never do. Moreover, whereas national studies of occupational prestige tend to emphasize consensus and harmony, studies of related occupations would naturally lead on to the study of status politics.

For all these reasons, we would suggest that investigations of these specific status orders would be a useful preliminary to the more difficult task of studying the status structures of entire societies.

APPENDIX
SOURCES FOR ORDERS OF PRECEDENCE INCLUDED IN TABLE 1

1. "Names of those elected and sworn in divers misteries of London for the government and instructions of the same." Pp. 232–4 in R. R. Sharpe (ed.), Calendar of Letters Books E. London: 1903.
2. "Funeral procession for King Henry V." P. 3 in R. R. Sharpe (ed.), Calendar of Letters Books K. London: 1911.
3. "Crafts to meet king on horseback upon his visit to the city." City Record, Journal 7, fo. 21b.
4. "Master and wardens to go Guildhall with armed men of crafts after flight of Edward IV." City Record, Journal 7, fo. 224.
5. "Order of going for every craft in the time of Sir Henry Colet, mayor." P. 66 in Charles Welch, History of the Worshipful Company of Pewterers of the City of London, Volume 1. London: Blades, East, and Blades, 1902.
6. "Order of procession for King Henry VIII's coronation." City Record, Journal 10, fo. 370b.
7. "Order of precedence for all goings and processions of the city companies." City Record, Repertory 3, fo. 66b.
8. "Companies' order of seating at the Mayor's feast." Pp. 190–2 in C. L. Kingsford (ed.), A Survey of London by John Stow, Reprinted from the Text of 1603. Oxford: Clarendon Press, 1908.
9. "Assessment of companies to pay for wheat for the poor." P. 135 in William Herbert, The Twelve Great Livery Companies of London, Volume 1. London: author, 1837.
10. "Provision of gunpowder." City Record, Journal 22, fo. 37b.
11. "Loan by city companies to provide ships, pinaces and 1200 men to join HM Navy." City Record, Journal 23, fo. 44b.
12. "Provision by companies of 10,000 qtrs. of corn for poor." City Record, Journal 30, fo. 61.
13. "Second report of the commissioners appointed to inquire into the municipal corporations in England and Wales. P. 249 in Parliamentary Papers, Vol XXV, Session 31 Jan–17 July, London, 1837.

REFERENCES

Balfour, M. Melville
1927 London Pride: Songs of the City. London: Homeland Association.

Barker, T. C.
1957 The Girdlers' Company: A Second History. London: private.

Blackham, R. J.
1931 London's Livery Companies. London: Sampson Low Marston.

Brentano, Lujo
1870 On the History and Development of Gilds and the Origin of Trade Unions. London: Trubner.

Clode, Charles Mathew
1875 Memorials of the Guild of Merchant Taylors' of the Fraternity of St. John The Baptist in the City of London and of its associated charities and institutions, compiled and selected by the Master of the Company for the year 1873–74. London: private.

Dumont, Louis
1970 Homo Hierarchicus: The Caste System and Its Implications. London: Paladin.

Girtin, Thomas
1958 The Golden Ram: A Narrative History of the Clothworkers' Company, 1528–1958, Second Edition. Aylesbury: private.

Hatt, Paul K.
1950 "Occupation and social stratification." American Journal of Sociology 45:539–43.

———
1906 The Guilds of Florence. London: Methuen.

Herbert, William
1837 The History of the Twelve Great Livery Companies of London, Volume One. London: author.

James, Margaret
1930 Social Problems and Policy During the Puritan Revolution, 1640–1660. London: Routledge.

Kellett, J. R.
1958 "The breakdown of gild and corporation control over the handicraft and retail trade in London." Economic History Review, Second Series: X,3.

Noorthouck, John
1773 A New History of London. London: Baldwin.

Pipes, Richard
1977 Russia under the Old Regime. London: Peregrine.

Reddaway, T. F.
1975 The Early History of the Goldsmiths' Company, 1327–1509. London: Edward Arnold.
Reynolds, H. T.
1971 Making Causal Inferences with Ordinal Data. Chapel Hill: University of North Carolina, Institute for Research in Social Science.
Treiman, Donald J.
1977 Occupational Prestige in Comparative Perspective. New York: Academic Press.
Unwin, George
1963 The Gilds and Companies of London, Fourth Edition. London: Frank Cass.
Wall, Cecil and H. Charles Cameron
1963 A History of the Worshipful Society of Apothecaries of London, Volume One, 1617–1815. Revised, annotated, and edited by E. Ashworth Underwood. London: Oxford University Press.
Wright, D. (ed.)
1877 History of Nepal. Trans. Munshi Shew Shunker Singh and Pandit Sri Gunand. Cambridge: Cambridge University Press.
Young, Sydney
1890 The Annals of the Barber Surgeons of London, compiled from their records and other sources. London: Blades, East, and Blades.

SOCIAL RESOURCES AND STRENGTH OF TIES: STRUCTURAL FACTORS IN OCCUPATIONAL STATUS ATTAINMENT*

NAN LIN
State University of New York at Albany
WALTER M. ENSEL
Albany Medical College
JOHN C. VAUGHN
National League for Nursing

For a class of social actions such as seeking a job, the socioeconomic standings of the contact (social resources) an individual uses will probably be very important in achieving a desired result. Drawing upon data from a sample of working males aged 21–64 in the metropolitan area of Albany-Troy-Schenectady, New York, we found that the job seeker's personal resources (initially his family background, but more importantly later his educational and occupational achievements) as well as his use of weak ties affect his ability to reach a contact of high status. The contact's status, in turn, has a strong and direct effect on the prestige of the attained job. As job experience increases, a person relies more on constructed rather than ascribed relations and the strong tie between his contact and the hiring firm becomes increasingly important.

The prevailing model of the status attainment process (Blau and Duncan, 1967; Duncan, Featherman, and Duncan, 1972; Sewell and Hauser, 1975; Jencks et al., 1972; Jencks and Rainwater, 1977; Featherman and Hauser, 1978) analyzes the demographic and distributive process of achievement and locates the occupational status which a person with certain personal resources can reasonably expect to attain. Criticism that the model lacks structural and relational concerns challenges the assumption that the labor market is essentially an open and competitive arena where specifications for a job and the necessary skills and competence are easily matched and where information about job and applicant availability is widely diffused (Rees and Schultz, 1970; U.S. Department of Labor, 1975). Researchers have been generally uncon-

* Direct all correspondence to Nan Lin, Department of Sociology, State University of New York at Albany, 1400 Washington Avenue, Albany, New York 12222.

An earlier version was read at the Annual Meetings of the American Sociological Association, New York, August, 1980. The study reported herein was partially supported by a research and development grant (No. 91-36-75-14) from the Manpower Administration, U.S. Department of Labor. We are grateful to Bonnie Erickson, Mark Granovetter, and Glenna Spitze for their comments on earlier drafts.

vinced that luck or chance largely accounts for the variation of occupational achievement which this prevailing model does not explain (Jencks, et al., 1972).

These criticisms do not reflect a lack of research into mobility patterns at the structural or institutional levels. Several studies, for example those by White (1970), Spilerman (1977), and Sorensen (1977), have investigated macro-level issues such as the movements of vacancies and occupants in a structure. The analysis of dual markets and, more recently, of segmentation of occupational and industrial groupings has attempted to show mobility restrictions and patterns embedded in the structure of the labor market (Gordon, 1972; Edwards, Reich, and Gordon, 1975). Stolzenberg (1978) has discussed how organizations affect status attainment. There have not, however, been sufficient studies of the process by which an individual gets a prestigious job or of the structural factors which help determine an individual's occupational status attainment.[1]

Our paper is not an attempt to construct a universal structural theory of status attainment; we focus on that portion of the process in which the job seeker evokes social relations. We do not even identify all the structurally relevant elements in this process. Rather, our effort is to demonstrate the fruitfulness of analyzing the status attainment process in terms of two particular structural factors, social resources and the strength of social ties.

NOTIONS OF STRENGTH OF TIES AND SOCIAL RESOURCES

Granovetter (1973:1361) has defined the strength of a tie as a "combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie." He sees weak ties as allowing a person to reach beyond his or her small, well-defined social circle in order to make connections with parts of the social structure not directly accessible to him or her. Weak ties serve as channels through which "ideas, influences and information socially distant from ego may reach him" (p. 1371). Granovetter speculated that the use of weak ties plays an important role in diffusing information and influence, providing opportunities for mobility, and helping the community to function. Boorman (1975) has demonstrated mathematically that under certain assumptions economic actors will maximize their benefits by maintaining weak ties and not strong ties.

Granovetter studied a random sample of professional, technical, and managerial personnel who had recently found new jobs through social contacts to determine how frequently the job seeker and the contact saw each other. He found that 16.7% of the job seekers saw their contact often at the time of the job search, 55.6% occasionally, and 27.8% rarely. Assuming frequency to indicate strength of tie, Granovetter concluded that job seekers used weak ties rather than strong ones. Further, he found those who had used weak ties were more satisfied with their new jobs than those who had used strong ties.

It is temping to go one step further and speculate that weak ties rather than strong ones are advantageous to individuals seeking higher status jobs. This is reasonable if we make certain assumptions, including the usual assumption that the overwhelming majority of job seekers prefer a more prestigious job to one with less prestige. We ignore the situation in which a job seeker must choose between two or more identically prestigious jobs. We may assume, in addition, that labor market information is not distributed uniformly in the social structure. The wider people cast their social nets, the more likely they are to gather information about a specific job. We also assume that the credibility of the referring person or agent

[1] Recognition of structural factors is evident in the status attainment literature. For example, Duncan, et al. (1968) demonstrated the reciprocal effect of friends' aspirations on ego's own occupational and educational aspirations, but their model did not include any status attainment measures. Sewell and Hauser (1975) examined interpersonal influence as social psychological factors in their study of status attainment among male Wisconsin high-school graduates. They found that certain factors (parents' encouragement, teachers' encouragement, and friends' plans to attend college) helped explain educational attainment, but not occupational attainment (pp. 94–6, 186).

influences an employer's decision to hire a particular person. It follows that a person is more likely to find the "right" or "influential" contact if he/she looks beyond his/her immediate social circle. These assumptions need not hold simultaneously, because some contacts will provide only information (knowledge about the precise job requirements and procedures for applying it) and others will exercise influence (ability to link the person to a particular segment of the labor market and enhance his/her chances of finding a job there). Nonetheless, these assumptions lead us to hypothesize that the strength of the tie used in job-seeking is negatively related to the occupational status one attains. In general, we should find that weak rather than strong ties help the job seeker find the desired job.

More recently, we (Lin, Vaughn, and Ensel, 1981) have focused on the social resources embedded in an individual's social network. Social resources are defined as "the wealth, status, power as well as social ties of those persons who are directly or indirectly linked to the individual." Lin (Lin, Dayton, and Greenwald, 1978) proposed that an individual's access to the social resources is instrumentally important. This proposal conceives of the social structure as comprising a network of persons whose positions are ranked according to certain normative honors and rewards, such as wealth, status, and power.[2] It further assumes that the structure has a pyramidal shape in terms of the accessibility and control of such honors and rewards.[3] A position nearer to the top of the structure has greater access to and control of honors and rewards not only because more honors and rewards are attached to the position intrinsically, but also because the higher position has greater access to positions elsewhere (primarily lower) in the rankings. In other words, a higher position commands a greater number of social resources.

The concept of social resources encompasses two components: social relations and the resources embedded in positions reached through such relations. The concept contrasts and complements the concept of personal resources as described in the social mobility and status attainment literature (Sorenson, 1977). While personal resources involve the individual's wealth, status, and power, social resources are embedded in the positions of contacts an individual reaches through his social network. These characteristics include but are broader than the reputational or prestigious characteristics as emphasized in the works of Laumann (1966) and Goode (1978). Our definition of resources is also consistent with Goode's four types of resources in the social-control processes: force, wealth, prestige, and friendship-love-affection (pp. 2–6). The last type delineates certain forms of social relations.

It would follow, then, that when seeking a job an individual will gain more by contracting someone upward in the hierarchical structure, who has, in other words, greater social resources. We hypothesize that reaching greater social resources is positively related to occupational status attainment.

Drawing upon data collected from a representative sample of males in the 21–64 age range among the noninstitutional civilian labor force, we (Lin, Vaughn and Ensel, 1981) identified those who had used social contacts in seeking first and current jobs. We then incorporated the contact's occupational status (Duncan's SEI score) into the basic Blau-Duncan status achievement model. The contact-status variable increased significantly the portion of the explained variance both for the first and current job statuses (about 10% in each case). Further, when we decomposed the effects, we found that the direct effects of

---

[2] The view that social structure consists of interlocking positions has gained support recently (White et al., 1976; Boorman and White, 1976; Burt, 1977). The social resources approach suggests that there should be some hierarchical order among clusters of positions, and its empirical demonstration awaits research efforts (see Burt, 1979, and Breiger, 1980, for approaches to flushing out such relationships).

[3] Violation of the pyramidal shape of a structure does not affect the general formulation. However, it is convenient to construct such an image to describe the panoramic view from the top toward the bottom of the structure and to indicate the relative number of occupants at each level of the structure. Sorensen's modelling of job earnings distribution (1977, Equations 1 and 2) assumes a similarly pyramidal inequality in society.

family background and education were much reduced and that much of the family-background effect was transmitted indirectly through the contact-status variable. We concluded, therefore, that "when one's social network is utilized for job seeking, his attained occupational status depends chiefly on his education and the social resources embedded in the network, which in turn depend largely on his ascribed status (family background)."

Strength of tie and social resources are, of course, related. They are concepts articulating an individual's interactions with structures of statuses and roles (Granovetter, 1979), and thus seem to be useful concepts in our attempt to bring social structure into an analysis of the status attainment process.

Conceptual Integration

The strength-of-ties literature has produced a hypothesis consistent with the well-known homophily (or like-me) principle (Homans, 1950; Laumann, 1966; Laumann and Senter, 1976; Verbrugge, 1979). This principle states that social interactions tend to take place among individuals with similar attributes. If frequent interaction indicates strong ties, then we may derive from the homophily principle the proposition that strong ties link persons of similar attributes. An extension of this would suggest that the heterophily principle (the linking of people with dissimilar attributes) should be more prominent among weak ties. Presumably, then, when seeking a job weak ties rather than strong ties will provide the seeker with a more extensive reach and hence with a greater likelihood of contacting people who possess job-related information and influence.

However, if the distribution of job information and influence is not random, but is part of the hierarchical structure of social resources, then the probability of an individual's gaining access to job information and influence increases when he or she makes contact with positions higher up in the structure. In other words, the heterophily principle may be ineffective if the positions reached are horizontal or lower in the structure relative to the person's initial status. We may call the tendency to contact positions at higher status levels in the structure the *prestige principle* (Laumann, 1966). It is the prestige principle, then, rather than the heterophily principle, that ought to operate if job seekers wish to maximize their chances of finding contacts who are sources of job information and influence. It is here that the linkage between the strength of ties and social resources occurs. If we assume that social resources, and hence job information and influence, increase at higher levels of the hierarchical structure, then an individual can most easily reach upward in a structure if he or she uses weak ties.

What we envision is a process whereby a job seeker reaches up to a contact as a source of the information or influence the job seeker needs to find a prestigious job. The job obtained is usually at a lower level than the ultimate source of information or influence. This process is illustrated in Figure 1. The first hypothesis states that for a given status level of the ego (E), the status of the contact reached (C or C′) is a function of the strength of the tie between the ego and the contact: weaker ties tend to reach contacts at higher levels (C rather than C′). A second hypothesis states that the status of the destination (D or D′)—the job obtained—is a direct function of the status level of the contact (C or C′) and an

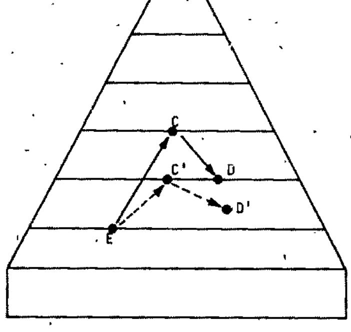


Figure 1. The Use of Social Resources in a Hierarchical Social Structure.

indirect function of the strength of the tie between the ego and the contact.

THE DATA AND ANALYSES

The data were from a modified random sample of males, 20–64 years of age, who resided in the tri-city area of Albany-Schenectady-Troy, New York, in the spring of 1975, and who were then or had been members of the United States civilian labor force. The total sample consisted of 399 respondents.

We asked each respondent whether any personal or impersonal channels were effective in helping him to get his first and current jobs. Three types of channels were used to find first and current jobs. For the first job, 57% of the respondents used personal contacts, 22% went directly to an employer, and 21% used formal channels. For the current job, 59% used personal contacts, 17% applied directly to an employer, and 24% used formal channels. To ascertain the effects of type of channel used on status of first and current jobs, we entered three dummy variables based on the above categories into a regression along with the basic status attainment variables. Only the use of formal channels showed a small positive effect on current job. While there was no evidence that the use of personal contacts had any significant effect on the attainment process, our conceptual model had led us to believe that the characteristics of the contacts used might have an effect on the occupational status attained.

We shall discuss further only those respondents who utilized social ties as channels in seeking their first or current jobs.

When a respondent mentioned a personal contact, we identified the contact's attributes. We measured social resources by the contact's occupational status (Duncan's SEI). Thus, there are two social resources measures: the occupational status of the contact used for the first job (O_{TW}) and the occupational status of the contact used for the current job (O_{TY}). We also obtained data on each contact's educational achievement, an alternative measure of social resources which we will discuss later.

We indexed the strength of the tie between the respondent and the contact as being either (1) weak, that is, acquaintances and indirect ties (e.g., friend's relatives, relative's friends), or (2) strong, that is, relatives, friends, and neighbors.[4] This variable was identified as S_{TW} for the relationship between the respondent and the contact relative to seeking the first job and as S_{TY} for that in seeking the current job. (In addition, we examined whether it was the respondent or the contact who initiated the interaction or whether the contact came about by chance. As it turned out, this variable did not have any significant effects, either direct or indirect, on the job the respondent attained.)

Effects of Strength of Tie and Social Resources on Attained Status

In the first phase of the data analysis we examined the interrelationships among the strength of tie between each respondent and the personal contact he utilized, the occupational status of the contact, and the attained occupational status of the respondent. We conducted separate analyses relative to the first and current jobs (see Table 1).

Comparing the contact's status with the occupational status attained by the respondent makes it evident that the former is significantly higher than the latter ($\overline{X}_{OTW} > \overline{X}_W$ and $\overline{X}_{OTY} > \overline{X}_Y$). This observation contradicts what the like-me principle would predict. If a person were best served in the social structure by ties with persons of similar status, as dictated by the like-me principle, we should have observed the respondent and his contact to have similar status. Instead, our results are consistent with the prestige principle: job seekers seem to reach upward when using contacts to get a job.

As expected, the zero-order correlations show significant relationships among

[4] We have here classified neighbors as strong ties, although we recognize that some researchers have called them weak ties (e.g., Wellman, 1979). However, given the fact that less than 3% of all individuals in our sample used neighbors to attain either first or current jobs (see Table 2), this classificatory difference does not appear very significant. For other possible criteria for measuring the strength of ties, see Granovetter, 1973.

Table 1. Means, Standard Deviations, and Zero-Order Correlations for the Strength of Tie, Contact Status, and Attained Status

	Seeking First Job				
	S_{TW}	O_{TW}	W	Mean	SD
Strength of tie to contact (S_{TW})	—	−.403	−.230	1.210	.408
Occupational status of contact (O_{TW})	202	—	.645	46.015	24.826
Attained occupational status (W)	204	205	—	36.795	24.028
	Seeking Current Job				
	S_{TY}	O_{TY}	Y	Mean	SD
Strength of tie to contact (S_{TY})	—	−.292	−.305	1.339	.475
Occupational status of contact (O_{TY})	170	—	.679	53.428	21.942
Attained occupational status (Y)	171	173	—	45.421	25.497

NOTE: All correlations at $p < .01$. The pair-wise analyses were based on the Ns shown in the lower off-diagonal cells.

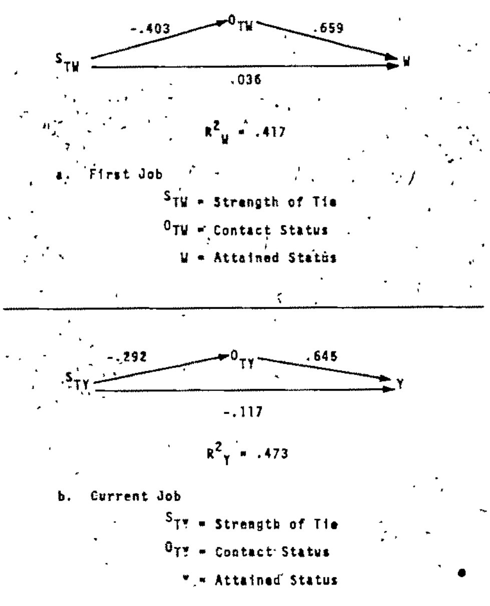


Figure 2. Effects of Strength of Tie and Contact Status on Attained Occupational Status

all three variables, suggesting the viability of both the strength-of-ties formulation and the social-resources formulation. The relative efficacy of the two formulations were examined by the extent to which the strength of the tie directly affected the attained occupational status. We simply regressed the attained occupational status on both the strength of tie and contact status variables and constructed path models. The results, in Figure 2, show that strength of tie has a weak (for the current job) or no (for the first job) direct effect on attained status. The social resources variable (contact status), on the other hand, retained its strong direct influence on attained status. The two independent variables explained 41% (for the first job) to 47% (for the current job) of the variation in attained status. The evidence suggests a concatenation in which the use of weak ties facilitates the reaching of higher status contacts, which in turn directly affects the attained occupational status.

The strength-of-ties perspective implies that weak ties lead to better social resources. From the social-resources perspective, however, the value of weak ties seems to depend on the job seeker's original position in the hierarchical structure. When his/her initial position in the structure is relatively low, only weak ties can provide access to social resources significantly higher up. But, when his/her initial position is relatively high, there is little reason to expect weak ties to have significantly greater advantage than strong ties. We hypothesize that the closer a job seeker's original position gets to the apex of the hierarchy, the more likely it is that strong ties will reach desirable social resources as well as weak ties. In fact, at the very top we should find strong ties to be more important, for weak ties will simply increase the likelihood of reaching downward in the structure.

We did not have the data to test the hypothesis regarding this limiting case; however, it was possible to test whether the weak tie became less significant as the job seeker's original status was higher in the hierarchical structure.

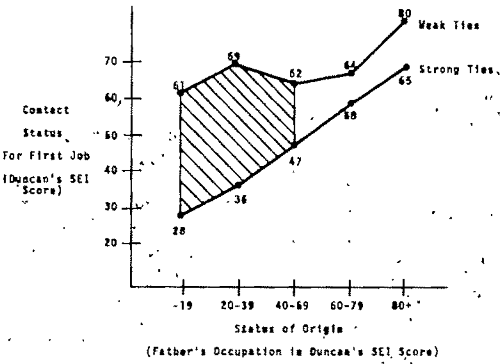


Figure 3. Relations Between the Status of Origin and the Contact Status Through Weak and Strong Ties (Shaded area indicates significant statistical differences.)

If we take father's occupational status to indicate the job seeker's original status, we can examine for each original status level whether a strong or a weak tie led to better social resources, as indexed by the status of the contact used for the first job. Figure 3 presents some data on this question. The benefit of using a weak tie seems constant; regardless of original status, the weak tie tends to lead to high-status contacts. When the original status is low, using a strong tie provided little benefit, but that benefit increases linearly as the original status goes up. In fact, toward the higher end of the original status level, the benefit of using a strong tie is almost as great as that of using a weak tie.[5] One must exercise caution in interpreting these trends, for the highest levels of original status may incur a ceiling effect. However, this caution itself is meaningful only if one adopts the social-resources perspective postulated earlier. Furthermore, what the data show is not that the use of weak ties declines in effectiveness, but rather that the use of strong ties becomes more helpful.[6]

Strength of the Tie between the Contact and the Firm

The strength-of-ties literature focuses on the relationship between the ego and the contact. It says little about the relationship between the contact and the destination—in the case of job seeking, the relationship between the contact utilized and the employing firm. A superficial application of the strength-of-tie variable would suggest that a weak tie between the contact and the firm would be more advantageous for the job seeker. This, of course, is false. From Granovetter's discussion (1973) it should be clear that the weak tie is useful to the ego because it extends his or her reach by being close to the desired destination and thus possessing information about and influence with the destination—in this case the hiring firm. Thus, a more appropriate hypothesis states that a weak tie between the job seeker and his or her contact tends to be related to a strong tie between the contact and the firm, which in turn is related to attaining a higher status job.

Similarly, the social-resources literature has not spelled out how the relationship between the ultimate contact and the hiring firm brings about the job seeker's being hired for the job. Again, interpreting the intent of the conceptual formulation, our hypothesis states that better social resources are related to a strong tie between the contact and the firm, which in turn is related to attaining a higher status job.

In the Albany survey each respondent who used a personal contact in job seeking

---

[5] The interactive effect of the status of origin and the strength of ties on the contact status can also be seen in the zero-order correlation between the simple product term and the contact status [$r(O_F \cdot S_{TW})(O_{TW}) = .518$]. Much of this effect of course is due to the association between the status of origin and the contact status ($r = .497$). This becomes apparent when the interaction term is expressed in terms of deviations from the means $(O_F - \overline{X}_{OF})(S_{TW} - \overline{X}_{STW})$. Its zero-order correlation with O_{TW} is .120. A regression analysis in which contact status is regressed on S_{TW}, O_F and $(O_F - \overline{X}_{OF})(S_{TW} - \overline{X}_{STW})$ reveals that the first two terms explained 34.5% of the variance and the interaction term contributed an additional 2.6%.

[6] It would be illuminating to use the attained status of the first job (W) as the ultimate dependent variable in the analysis presented in Figure 3. However, a four-way breakdown (W by O_F, S_{TW}, and O_{TW}) results in three empty cells and extremely small frequencies (between 2 and 22) in all except one remaining cell. Statistical tests could not be conducted. A similar plot for the current job was not done, because we could not justify using W as the status of origin and ignoring the job changes that had taken place between W and Y. A treatment of the model taking into account the individual's job history is beyond the scope of the paper.

was asked to characterize the relationship between his contact and the hiring company. Responses were put into two groups: (1) weak—the contact knew someone in the company or someone who had an acquaintance in the company—and (2) strong—the contact owned the company or worked for the company.[7] This variable is hereafter identified as F_{TW} relative to the first job, and F_{TY} relative to the current job.

While conceptually we can compare the strength of the tie between the respondent and the contact and between the contact and the firm, the measures we used for the two relations in this study differed. That is, we measured the strength of tie between respondent and contact by their relationship (i.e., whether they were relatives, friends, neighbors, acquaintances, etc.). We measured the strength of the tie between the contact and the firm by whether the contact owned the firm, was employed there, or was altogether unconnected with it. The latter measure is admittedly crude. We hope better measures will be devised in future studies.

Again, we constructed structural equations models for the occupational attainment process for the first and current jobs. These results are shown in Figure 4 as path models. The strength of the tie between the contact and the firm does not appear to be an important variable in finding a prestigious first job. It does not contribute significantly to the attained status of the first job, nor is it much affected, in a linear fashion, by the strength of the tie between the respondent and the contact or by the contact's status.

As for the current job, a strong tie between the contact and the firm shows a moderate direct effect on the attained status. The strength of the tie between the contact and the firm is in turn directly affected by the contact's status, but not by the strength of the tie between the job seeker and the contact.

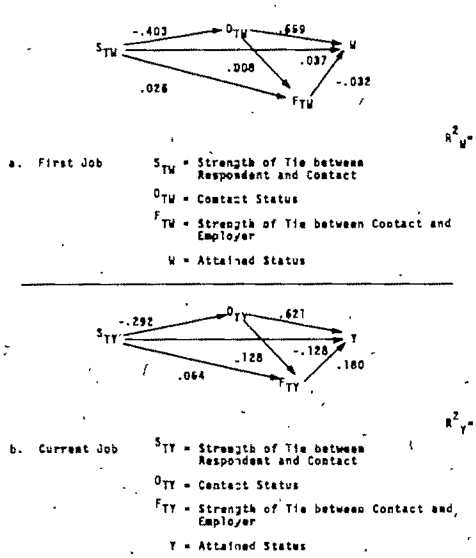


Figure 4. Effects on Attained Occupational Status of Strength of Tie Between Respondent and Contact, Contact Status, and Strength of Tie Between Contact and Firm

The anomalous finding that the strength of the tie between the contact and the firm does not contribute to explaining the attained status of the first job requires further deliberation and analysis.

We wondered whether the job seeker's personal resources (e.g., father's occupation and education, and job seeker's own education) at different phases of his occupational career affected the job-seeking process and the contacts and resources evoked. One speculation was that when entering the occupational structure (seeking the first job), the job seeker's personal resources would greatly affect the social relations he could utilize. Thus, parental characteristics should play an important role in reaching the "right" contacts and resources, not necessarily because these contacts and resources have direct links with possible employers, but because they represent the farthest that the job seeker can reach. Once a person begins working, he can make more of his own contacts, many of which are work-related. One ceases to rely largely

[7] For the STW group, 26.8% of the contacts knew an employee in the company or knew someone who had an acquaintance in the company; 67.3% were an employee in the company; and 5.9% owned the company. For the STY group, 21.5% knew an employee or had an acquaintance who knew an employee; 74.2% were employed by the company; and 4.3% owned the company.

A separate analysis showed no differential effects on status attainment whether the contacts were owners or employees, so we collapsed the two categories.

Table 2. Types of Contacts Used for First and Current Jobs

Nature of the Tie Between Respondent and Contact	First Job (N = 185) %	Current Job (N = 219) %
Strong ties		
Neighbor	2.8	2.8
Relative	41.1	26.1
Friend	35.5	37.2
Weak ties		
Acquaintances	15.4	32.8
Other indirect ties (relative's friend, friend's relative)	5.1	1.2

on prescribed (both consanguineous and affinial) relations and comes to rely on one's own constructed relations. As a person becomes more embedded in a network of constructed relations, contacts tied to a potential employer increase in significance.

Table 2 shows the relative frequency of various types of contacts used for obtaining first and current jobs. For the first job, 41% of the respondents using personal contacts went to relatives, 35% to friends, and 21% to acquaintances and indirect links (e.g., relative's friend or friend's relative). For the current job, only 26% used relatives, friends accounted for 37%—about the same as for first jobs—while use of acquaintances and indirect links increased to 34%. Thus, there is some evidence that constructed ties gradually replace prescribed ties (relatives) as one progresses in an occupational career. Granovetter (1974:42) obtained comparable data. The direct impact of personal background on one's status appears to decrease as one's occupational career develops.

To examine in detail such processes would require data beyond what was available to us. However, we were able to incorporate some personal resources variables into the models and explore their differential effects on the social resources factors as well as on the attained status of jobs.

Differential Impacts of Personal and Social Resources on Status Attainment

We examined the variables in the basic Blau-Duncan model, as well as the three structural factors (the strength of the tie between the respondent and the contact, the contact's status, and the strength of the tie between the contact and the firm) for their effects on the job-seeking process. However, this is not exactly the same as expanding the Blau-Duncan model, since we analyzed only data from respondents who used personal contacts in seeking first and/or current jobs. We wanted to examine further certain hypotheses derived from the social-resources and strength-of-ties literature. Nonetheless, the analysis did allow us to determine after we had taken personal resources variables into account whether the effects of the structural factors on job seeking were spurious.

The analysis entered the variables in the following causal sequence: (1) the father's education (E_F) and occupational status (O_F) as the exogenous variables; (2) the job seeker's education (E); (3) the strength of the tie between the job seeker and the contact; (4) the contact's occupational status; (5) the strength of the tie between the contact and the firm; and (6) the job seeker's attained status. For the current job, we also entered the occupational status of the first job as an independent variable. Again, we conducted separate analyses relative to the first and current jobs. The results appear in Tables 3 and 4 respectively.

For the first job, personal resources (father's occupation and job seeker's education) have a significant effect on the status of the contact reached, but not on the strength of the tie evoked. The strength of the tie between the job seeker and the contact, however, affects the contact's status as well. Personal resources and use of weak ties appear to have independent effects on reaching higher status contacts. However, neither the strength of the tie between the job seeker and the contact nor that between the contact and the firm directly affect the prestige of the first job. When the job seeker's personal resources are taken into account, the contact's status shows the strongest direct effect among all independent variables on getting a prestigious first job.

In gaining the current job, the status of

Table 3. Ordinary Least Squares Estimates for First-Job Status Attainment

Dependent Variable	Independent Variables							Error of Estimate
	E_F	O_F	E	S_{TW}	O_{TW}	F_{TW}	a	
	Metric Coefficients							
Education	.321	.046						
(E)	(.063)	(.011)	—	—	—	—	7.76	2.58
Tie to respondent	.003	−.002	−.019					
(S_{TW})	(.010)	(.002)	(.011)	—	—	—	.92	.40
Contact status	.407	.255	2.05	−17.58				
(O_{TW})	(.507)	(.085)	(.558)	(3.57)	—	—	−14.45	19.36
Tie to firm	.001	−.001	.001	.027	.000			
(F_{TW})	(.001)	(.002)	(.001)	(.088)	(.002)	—	1.23	.45
First-job status	.495	.118	1.98	2.01	.411	−1.30		
(W)	(.437)	(.075)	(.497)	(3.27)	(.064)	(2.75)	−15.55	16.66
	Standardized Coefficients						R^2	
(E)	.382	.320	—	—	—	—	.409	
(S_{TW})	.033	.132	−.152	—	—	—	.053	
(O_{TW})	.065	.240	.276	−.289	—	—	.405	
(F_{TW})	.001	−.034	.001	.025	.026	—	.001	
(W)	.082	.115	.276	.034	.425	−.024	.535	

NOTE: Standard errors appear in parentheses below their corresponding metric coefficients.

the first job, rather than other personal resources, has the strongest direct effect on the strength of the tie between the job seeker and the contact and on the contact's status.

The job seeker's educational achievement and the status of his contact have a modest effect on the strength of the tie between the contact and the firm. The contact's status and his strong tie to the firm, as well as the seeker's education and first-job status, directly aid in acquiring a prestigious current job. These results further confirm that family background gradually ceases to affect directly the choice of contact or the outcome of the job search, while increasingly work status (as reflected in first-job status) does. The contact's status and his tie to the firm directly affect acquisition of a prestigious

Table 4. Ordinary Least Squares Estimates for Current-Job Status Attainment

Dependent Variable	Independent Variables								Error of Estimate
	E_F	O_F	E	W	S_{TY}	O_{TY}	F_{TY}	a	
	Metric Coefficients								
Education	.254	.047							
(E)	(.065)	(.010)	—	—	—	—	—	8.59	2.44
First job	.362	.217	4.01						
(W)	(.505)	(.081)	(.588)	—	—	—	—	−25.01	18.19
Tie to respondent	.011	−.002	−.005	−.005					
(S_{TY})	(.013)	(.002)	(.017)	(.002)	—	—	—	1.14	.46
Contact status	.873	.088	.925	.244	−7.21				
(O_{TY})	(.499)	(.082)	(.657)	(.079)	(3.07)	—	—	10.65	17.56
Tie to firm	−.003	−.001	.024	−.002	.051	.002			
(F_{TY})	(.012)	(.002)	(.016)	(.002)	(.075)	(.002)	—	1.46	.41
Current job status	.001	−.007	3.35	.129	−4.78	.426	9.39		
(Y)	(.001)	(.063)	(.558)	(.070)	(2.67)	(.069)	(2.96)	−20.00	14.70
	Standardized Coefficients							R^2	
(E)	.321	.372	—	—	—	—	—	.402	
(W)	.058	.217	.502	—	—	—	—	.481	
(S_{TY})	−.093	−.094	−.031	−.242	—	—	—	.079	
(O_{TY})	.158	.100	.132	.278	−.156	—	—	.380	
(F_{TY})	−.030	−.029	.185	−.112	.059	.123	—	.032	
(Y)	.001	−.006	.412	.127	−.089	.366	.152	.681	

NOTE: Standard errors appear in parentheses below their metric coefficients.

job, even after the personal resources variables have been taken into account. We should also note that both the structural factors and the personal resource variables explained 67% of the variation in the attained status.

While the analyses only involved job seekers who used contacts in seeking first and current jobs, the fact that the basic Blau-Duncan model was verified (see the relative direct contributions of job seeker's education and father's occupation on the first-job status in Table 3 and those of job seeker's education and first job on the current-job status in Table 4) increases our confidence that our three structural factors do contribute to the status attainment process.

DISCUSSION

The picture which emerges from these preliminary conceptual and empirical deliberations shows that the social resources an individual job-seeker evokes have a significant association with the status of the job he attains. Whether the job seeker can utilize a high-status contact depends on both his personal resources (initially his family background but eventually his educational and occupational achievements) and his use of weak ties. As the job seeker progressively comes to rely on his constructed relations rather than his ascribed relations, his contact's strong tie with the hiring firm becomes increasingly important.

Of course, the strength of ties and social resources are not the only important structural factors in the status attainment process, nor are our measurements anything but rudimentary. But our data did permit us to measure social resources in terms of the contact's educational status.

The zero-order correlation between the contact's educational status and respondent's first and current occupational status (.519 and .559, respectively) was comparable to that between the contact's occupational status and the respondent's first and current occupational status (.645 and .679, respectively). When we substituted the contact's educational status for his occupational attainment, the variables in the model explained almost 50% of the variance (47.8%) in first job and more than 60% (60.2%) of the variance in current job.

We collected data on both first-job income and current-job income. These could be used as alternate measures of job prestige. We did not have enough confidence in the first-job income data because it required overtime adjustments and because the respondents could not always remember the figures accurately. We decided to focus on the current-job income.

When we examined current-job income as the dependent variable, the analysis showed that both the strength of the respondent's tie to the contact and the contact's occupational status related significantly to current-job income (r = −.264 and .438, respectively). In addition, the mean income of respondents using weak ties was some $2,500 more than those using strong ties. When we entered the contact's occupational status and the strength of the tie between the respondent and the contact as well as the strength of the tie between the contact and the firm into the regression equation along with the basic status attainment variables, the model explained almost 31% of the variance in current income. Compared to other models of income attainment (Sewell and Hauser, 1975), this increases by more than 15% the variance explained in current-job income.

This, it appears that alternative measures of both the contact's characteristics and the job seeker's socioeconomic status demonstrated patterns similar to those discussed earlier.

Another explanation (the like-me principle, Laumann, 1966) suggests that a higher social position has advantage because it connects to many other positions on the same level. This should mean that a person's status would be similar, in general, to that of his contact—that the higher a person's status, the higher the status of his or her contact. But, we did not find this to be the case (see the differential contact and attained job statuses in Table 1).

An earlier discussion (Lin, Vaughn, and Ensel, 1981) pointed out that we need to determine whether the structural variables incorporated in the model could be proxy variables for the criterion variables. For example, we may ask whether the contact's occupational status merely reflects

the occupational status the job seeker desires (the dependent variable) or whether it is a causal factor. This is a difficult task empirically. However, we conducted a preliminary examination of the proxy hypothesis with the available data and found no evidence of its presence.

If a proxy variable is present (call it Y′) in terms of the finally achieved occupational status (current job, Y), Y′ and Y should have two empirical properties, at least as far as internal validity of the attempted causal modelling is concerned:

(1) Y′ and Y should have approximately the same status level. This is a strong statement of the proxy hypothesis. However, our data show the contact status (O_{TY}) to be much higher than the achieved status (Y). This can be seen in Table 3 (the means are statistically different when the standard deviations are used to compute the standard errors).

(2) Y′ and Y should have similar causal patterns in terms of their relationships with the independent variables. Negation of statement 1 does not necessarily reject the proxy hypothesis, since the proxy variable may involve a different type of measure and be subject to a different amount of error. Thus, a more definitive test is to examine the patterns of relationships between the dependent variable and its proxy and the independent variables. In our case, we needed to examine the two estimation equations.

On the right side of each equation the first two terms are, of course, the father's education and occupation; the third term is the job seeker's education; and the last term is the job seeker's first job—the usual status attainment causal formation. The proxy hypothesis does not insist that the parallel Bs have similar magnitude, since these independent variables should have greater effect on Y than on its proxy variable Y′. However, the patterns among the Bs must be similar to those among the B′s. The results of our data appear in Table 5. The last two coefficients (for E and W) are in reversed magnitudes as far as Y and Y′ are concerned.

Thus our data do not support the proxy hypothesis as reflected in statements 1 and 2.

That the characteristics of one's contact in an instrumental action are empirically distinctive from the characteristics of one's desired destination rather than their proxies was also confirmed in a smallworld study (Lin, Dayton, and Greenwald, 1978). In chains which eventually reached a target person, the intermediaries tended to have weak-tie relations with each other and to have higher status and prestige than both the starters and the target person. Since the smallworld study involved actual individuals in an actual chain process, it further validates the findings discussed in this paper.

Table 5. Test of Proxy Status Hypothesis (Y = Current Occupational Status, Y′ = Contact Status)

Dependent Variable	Estimated Coefficients			
	E_F	O_F	E	W
Y	.048	.034	.494	.250
Y′	.144	.115	.137	.316

Within conceptual and measurement confines discussed above, we have gained some insight concerning certain linkages between individuals and the social structure. To the extent that the hierarchical structure of society dictates the networks in which individuals are embedded, our discussion and analysis point to the potential theoretical usefulness of social resources and strength of ties as structural factors. For a class of social actions identifiable as instrumental for the individual involved (in this case, seeking a job), resources reached through his or her social network probably have much to say about the outcome of the action (in this case, the job attained).

Nor does this emphasis deny the part that personal resources contribute to instrumental actions. The preliminary evidence suggests an intricate interplay between personal resources and social resources. An individual's personal (especially prescribed) resources may greatly affect, at least initially, what social resources are available to him or her. As social resources are accumulated, however, these become more directly important than personal resources in further elaborating social networks and, therefore, social resources. In time, the indi-

vidual's accumulated social resources become the personal (and prescribed) resources for the next generation of actors.

REFERENCES

Blau, Peter M. and O. D. Duncan
1967 The American Occupational Structure. New York: Wiley.

Boorman, Scott A.
1975 "A combinational optimization model for transmission of job information through contact networks." Bell Journal of Economics, 6:216–49.

Boorman, Scott A. and Harrison C. White.
1976 "Social structure from multiple networks II: role structures." American Journal of Sociology 81:1384–446.

Breiger, Ronald L.
1980 "The social class structure of occupational mobility." Mimeographed paper.

Burt, Ronald S.
1977 "Positions in multiple network systems, part one: a general conception of stratification and prestige in a system of actors cast as a social typology." Social Forces 56: September.
1979 "Studying the status/role-set duality by mass survey: ersatz network positions and structural equivalence in large populations." Working Paper 27, Survey Research Center, University of California at Berkeley.

Duncan, Otis D., David L. Featherman, and Beverly Duncan
1972 Socioeconomic Background and Achievement. New York: Seminar Press.

Duncan, Otis D., A. O. Haller, and A. Portes
1968 "Peer influences on aspirations: a reinterpretation." American Journal of Sociology 74:119–37.

Edwards, Richard C., Michael Reich, and David M. Gordon
1975 Labor Market Segmentation. Lexington, Mass.: Heath.

Featherman, David and Robert Hauser
1978 Opportunity and Change. New York: Academic Press.

Goode, William
1978 The Celebration of Heroes. Berkeley: University of California Press.

Gordon, David M.
1972 Theories of Poverty and Underemployment. Lexington, Mass.: Heath.

Granovetter, Mark
1973 "The strength of weak ties." American Journal of Sociology 78:1360–80.
1974 Getting a Job. Cambridge, Mass.: Harvard University Press.
1979 "Toward a sociological theory of income differences." Paper read at the Annual Meetings of the American Sociological Association, Boston, Mass., August.

Homans, George C.
1950 The Human Group. New York: Harcourt, Brace, Jovanovich.

Jencks, Christopher, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephen Michelson
1972 Inequality. New York: Harper and Row.

Jencks, Christopher and Lee Rainwater
1977 The Effects of Family Background, Test Scores, Personality Traits, and Education on Economic Success. Report prepared for the National Institute of Education and the Employment and Training Administration of the U.S. Department of Labor.

Laumann, Edward O.
1966 Prestige and Association in an Urban Community. Indianapolis: Bobbs Merrill.

Laumann, Edward O. and R. Senter
1976 "Subjective social distance, occupational stratification and forms of status and class consciousness: a cross-national replication and extension." American Journal of Sociology 81:1304–38.

Lin, Nan, Paul Dayton, and Peter Greenwald
1978 "Analyzing the instrumental uses of social relations in the context of social structure." Sociological Methods and Research 7:149–66.

Lin, Nan, John C. Vaughn, and Walter M. Ensel
1981 "Social resources and occupational status attainment." Social Forces (June).

Rees, A. and G. P. Schultz
1970 Workers and Wages in an Urban Labor Market. Chicago: University of Chicago Press.

Sewell, William and Robert M. Hauser
1975 Education, Occupation and Earnings: Achievement in the Early Career. New York: Academic Press.

Sorensen, Aage B.
1977 "The structure of inequality and the process of attainment." American Sociological Review 42:965–78.

Spilerman, Seymour
1977 "Careers, labor market structure, and socioeconomic achievement." American Journal of Sociology 83:551–83.

Stolzenberg, Ross M.
1978 "Bringing the boss back in." American Sociological Review 43:813–28.

U.S. Department of Labor
1975 Job Seeking Methods Used by American Workers. Bureau of Labor Statistics Bulletin, no. 1886.

Verbrugge, Lois
1979 "The structure of adult friendship choices." Social Forces 56:576–97.

Wellman, Barry
1979 "The community question: the intimate networks of East Yorkers." American Journal of Sociology 84:1201–32.

White, Harrison C.
1970 Chains of Opportunity. Cambridge, Mass.: Harvard University Press.

White, Harrison C., S. A. Boorman, and R. L. Breiger
1976 "Social structure from multiple networks I: blockmodels of roles and positions." American Journal of Sociology 81:730–80.

TRENDS IN OCCUPATIONAL MOBILITY IN CANADA AND THE UNITED STATES: A COMPARISON*

HUGH A. MCROBERTS
KEVIN SELBEE
Carleton University, Ottawa

Using data from the Canadian Mobility Study and from the OCGI study, we compare occupational mobility in Canada and the United States for males who were 16 years of age between 1924 and 1953. The data are coded into six comparable occupational groupings and the problems involved in doing this are discussed. The analysis of trends is carried out both within and between countries. Using log-linear models, we find that there is no trend in mobility in either country and no difference in mobility between the two countries.

In their recent book, Hauser and Featherman (1977:15) state,

> We believe that industrial societies can be shown not to have the same rates of *observed* mobility (cf. Broom and Jones, 1969a, 1969b). However, there is reason to suppose that they may have similar patterns of *circulation* mobility. For two societies [the United States and England and Wales] where adequate longitudinal data are available, it seems that, once structural mobility has been taken into account, circulation mobility has been nearly constant over time.

Although this statement is a reasonable summary given the available evidence, that evidence is by no means conclusive. The earlier intersocietal comparisons of Lipset and Bendix (1959), S. M. Miller (1960), and Broom and Jones (1969a) are based upon the crudest of occupational classifications. Moreover, their methodological procedures do not wholly disentangle the effects of structural and circulation mobility. The other evidence offered in support of Hauser and Featherman's position is the rather more substantial United States–Australia comparison, which employs more detailed occupational categories than the earlier studies and uses a proven methodology (Hauser and Featherman, 1977:9–16).

The evidence for unchanging circulation mobility within societies is not much stronger. While studies by Hauser et al. (1975) and Hope (1974) are both of high quality, they apply to only two countries—the United States and England, respectively. Given the incompleteness of this evidential base, Hauser and Featherman's statement should be regarded as a set of hypotheses to be tested. Our purpose is to test them for Canada and the United States.

* Direct all correspondence to Hugh A. McRoberts or Kevin Selbee, Department of Sociology and Anthropology, Carleton University, Ottawa, Canada K1S 5B6.

This is a revised version of a paper presented to the Research Committee on Stratification and Mobility, International Sociological Association 9th World Congress of Sociology, Uppsala, Sweden, August 14–19, 1978. The research reported in this paper is part of the Canadian National Mobility Study, a collaborative project involving Monica Boyd, John Goyder, Frank E. Jones, Hugh A. McRoberts, Peter C. Pineo, and John Porter. This research is funded by the Canada Council. We would like to thank our colleagues Monica Boyd, Judah Matras and John Myles for their assistance, both in discussions and through comments on an earlier draft of this paper. Finally, we would like to thank Professor Robert M. Hauser and two anonymous reviewers for their valuable comments. We are, of course, responsible for errors and omissions which remain.

The data from *Occupational Changes in a Generation, 1962* (OCGI) utilized in this paper were made available by the Data and Program Library Service at the University of Wisconsin–Madison, through the Carleton University Social Science Data Archives. The data for OCGI were originally collected by the U.S. Bureau of the Census under grants from the National Science Foundation to Peter M. Blau and Otis Dudley Duncan of the University of Chicago. The data were reprocessed and released in unit record form by the Bureau of the Census under NSF grant Number GI-31604X to David L. Featherman and Robert M. Hauser of the University of Wisconsin–Madison. Neither the original collectors of the data, nor Messrs. Featherman and Hauser, nor DPLS bear any responsibility for the analysis or interpretations presented here.

These hypotheses, implied by Hauser and Featherman, can be stated as follows:

(1) While in an industrial society the pattern of observed mobility may vary with respect to time, the pattern of circulation mobility will be invariant with respect to time. (We will define circulation mobility in operational terms below.)
(2) While any pair of industrialized societies may have different patterns of observed mobility, they will have similar patterns of circulation mobility.

A third hypothesis is based on Hypotheses 1 and 2:

(3) In any pair of industrial societies, circulation mobility varies with respect to neither time nor country.

Given the research upon which they are based, these hypotheses should apply to the period from roughly the end of World War I to 1980.

Existing literature suggests that a comparison of circulation mobility in Canada and the United States will not support these hypotheses. The characterization of Canada as a closed society with respect to mobility (compared to the U.S.) is taken not so much as a hypothesis but rather as an established fact to be explained (Naegele, 1961; Lipset, 1963; Porter, 1965). Naegele and Lipset, working from a functionalist framework, attribute this difference to the nature of the "Canadian compromise." Canada is seen by them as a country which is neither British nor American, but which has attributes of each embedded in its value structure. In terms of the stratification values of elitism vs. egalitarianism, ascription vs. achievement, and particularism vs. universalism, their conclusion is that Canada lies somewhere between the United States and Britain; Canada is more open than Britain and less open than the United States. Treiman and Terrell's (1975) comparison of status attainment in Great Britain and the United States lends some support to the Naegele-Lipset thesis in showing that there is, as predicted, somewhat less mobility in Great Britain than in the United States.

In his discussion of stratification in Canada, Porter (1965) follows a different line of argument. While he makes some reference to value differences, he bases much of his explanation on the lack of openness of Canadian educational institutions. Specifically, he levels two charges at Canadian educational policy: that it has failed to supply the highly qualified manpower demanded by structural changes in both the economy and the labor force, and that it has remained elitist and closed to opportunity. Porter assumes implicitly that the relationship between educational attainment and occupational position is a constant (cf. Porter, 1965:166–7). He then argues that if educational opportunity is more closely related to family background in Canada than in other countries, especially the United States, the result will be that Canadians will also experience less mobility (Porter, 1965:49–59 and 163–8).

In short, those who have given some attention to the question of differences in mobility between Canada and the United States would argue, albeit for different reasons, that such differences do exist. In particular, they argue that we should expect to find that Canadians experience less mobility than Americans.

When we turn to the consideration of *intra*country variations in mobility, there is more direct evidence. In the American case, the initial findings of Hauser et al. (1975) of invariance with respect to time in circulation mobility have found substantial support in subsequent work.[1] Of particular interest are the findings of Featherman and Hauser (1978), who employ in part of their most recent analysis a five-fold set of occupational strata which is very close in concept to the six-fold breakdown which we will employ. In this extended analysis they do find some small temporal fluctuations in the patterns of mobility associated with particular occupational strata. However, even given these they conclude,

> In summary our reanalysis of the 1962 and 1973 classifications of mobility from father's occupation to son's first occupation reconfirms our earlier conclusions that intergenerational mobility chances were virtually constant for American men entering the labor

[1] The reanalysis of the British data presented by Goldthorpe et al. (1980) also confirms Hope's earlier analysis (1974) and gives results which are consistent with our hypotheses.

> market from about 1920 to 1970, while observed changes in mobility rates were induced by changing distributions of occupational origins and destinations. (1978:208)

In the Canadian case, there is some reason, based on two previous studies, to believe that there may well be some temporal variability in the rates of circulation mobility. Using two sets of data gathered in the Province of Quebec 10 years apart, Dofny and Garon-Audy (1969:285) find some evidence of increasing mobility, especially in the case of Francophones. However, using indices of association and percentages, they are not able to focus sharply on the patterns of pure circulation mobility.

In addition to the work by Dofny and Garon-Audy, McRoberts (1975) carried out an analysis of trend using data from the Canadian Mobility Study. Using log-linear models, he found that there was a slight—but discernable—trend, and further that the impacts on mobility of both the Great Depression and the Second World War were clearly identifiable in his data (1975:104–26). However, McRoberts was attempting to look at status mobility using "classes" based on analytic cutting points on the Blishen-McRoberts S.E.I. (1976). There is no necessity for an occupationally based analysis to give the same results.

Thus, there is some reason to expect that we will find intercountry variations in the pattern of circulation mobility as well as intracountry variations over time.

THE DATA

The data for this study are drawn from two sources: the Canadian Mobility Study (CMS) and the Occupational Change in a Generation study, carried out by Blau and Duncan (OCGI). It is not wholly an accident that the two studies are very similar in their design; the CMS was conceived in part as a replication of OCGI, and both were supplements to their countries' labor-force surveys. In design and sample, these surveys are also very similar, again not surprisingly, as the two agencies involved—Statistics Canada and the Bureau of the Census—consult frequently on such undertakings. (For more detailed examination of the similarities and differences between the two studies, see Boyd and McRoberts, 1974, and Blau and Duncan, 1967.) The methodologies of design, sampling, instrument construction, and data gathering are sufficiently similar that we feel that it is most unlikely that any bias in the findings is due to these factors.[2]

The populations to be compared will be, for Canada, native-born males aged 36–65 who would have been aged 25–54 in 1962, and, for the United States, nonblack males aged 25–54 in 1962. In Canada we decided to eliminate non-native-born males, as they constitute a significant element (15.3%) of the population in 1971 and other research suggests that they have somewhat different mobility experiences from those who are native born (Boyd and Porter, 1977). In the American case, we have excluded blacks because research has consistently shown that this group has a distinctively different pattern of achievement from American nonblacks. We have not excluded the foreign-born, both because this group is small (just over 7% in the OCGI sample) and because research has shown that nativity is not an important attainment factor in the United States (Featherman and Hauser, 1978:475).

OCCUPATIONAL CODING

In a comparative study of the type we are doing, the degree to which we can establish a satisfactory match between the oc-

[2] The Canadian researchers were well aware of the problems encountered by Blau and Duncan in the OCGI study with respect to first-job responses. While they were at some pains to avoid this problem in the design, there is some evidence that the respondents were not as cooperative as they might have been and gave occupations which were held prior to the completion of their schooling. In general, the effect of this is believed to have been relatively small, especially for native-born males. Thus, our results are more like those for OCGII than OCGI with respect to this variable. This means, however, that comparisons with the margins of the Canadian and American matrices are likely to be affected by the measurement of first job. However, since Featherman and Hauser show that the interactions between fathers and sons seem to be unaffected by this (1978:200–8), our analysis ought not to be seriously affected.

cupational codes in the two data sets is vital. Were it possible to return to the original instruments and code the responses to a common scheme, this would have been an issue of little consequence. However, at this time such an activity is not possible. As a result, we were forced to find a way of matching occupations using only the occupational codings employed in the data sets. In the Canadian data we had two items of information to work with, the 1971 Blishen-McRoberts S.E.I. (1976) and a sixteen-fold occupational classification, devised by Pineo, Porter, and McRoberts (PPM) (1977). The former was of little value to us, as the identification of unique occupational titles was virtually impossible due to rounding of the scores to two digits. In the American data we were more fortunate, as we had the 3 digit 1960 Census occupational codes available to us. We decided on a procedure which involved mapping the American codes onto the Canadian, Pineo et al., sixteen-fold occupational classification.

The matching procedure involved a number of problems described in detail in our appendices. We had originally hoped to preserve the level of detail in the sixteen-fold classification in our matching procedure. However, due to the differences in the way in which the Canadian and American classifications were constructed, we had to do some collapsing of the occupational categories, as shown in Table 1. Table 2 presents the distributions using these categories for respondent's first full-time job by cohort and by country.

Table 1. Relationship Between the Sixteen-Fold Occupational Classification[a] and the Six-Fold Collapsed Classification

1. Professionals and semiprofessionals
   - Self-employed professionals
   - Employed professionals
   - Semiprofessionals
   - Technicians
2. Managers and proprietors
   - High level managers
   - Middle managers
   - Supervisors
   - Skilled white collar
3. White collar workers
   - Semiskilled white collar
   - Unskilled white collar
4. Upper blue collar
   - Foremen
   - Skilled trades and crafts
5. Lower blue collar
   - Unskilled laborers
6. Farm and farm labor
   - Farmers
   - Farm laborers

[a] Pineo, Porter, and McRoberts (1977).

Table 2. First Full-Time Job by Cohort and Country

	Year Respondent Was Aged 16						
Canada	1924–28	1929–33	1934–38	1939–43	1944–48	1949–53	Total
Professionals and semiprofessionals	6.9	7.1	8.0	8.6	11.0	14.2	9.7
Managers and proprietors	7.9	7.6	6.0	7.4	7.2	8.0	7.3
White collar	14.7	15.0	14.7	14.2	15.0	15.6	14.9
Upper blue collar	33.0	30.6	35.1	35.5	34.9	32.4	33.8
Lower blue collar	12.2	17.1	16.9	19.5	19.9	21.5	18.6
Farm	25.3	22.6	19.3	14.8	11.9	8.3	15.7
Total	100.0	100.0	100.0	100.0	99.9	100.0	100.0
N	460	831	992	1062	1201	1133	5679
United States							
Professionals and semiprofessionals	8.4	6.0	8.1	8.7	12.0	13.7	9.4
Managers and proprietors	3.9	3.9	3.8	5.2	6.2	4.9	4.7
White collar	20.5	21.6	20.6	18.8	18.1	18.9	19.8
Upper blue collar	26.4	26.5	30.2	32.7	31.2	29.7	29.6
Lower blue collar	19.4	22.5	21.2	18.6	20.5	22.2	20.7
Farm	21.4	19.4	16.1	15.9	12.0	10.6	15.8
Total	100.0	99.9	100.0	99.9	100.0	100.0	100.0
N	1824	2054	2142	2260	2064	1825	12,169

The most serious problem we encountered, which was not wholly overcome, concerned the assignment of those who were proprietors. In the American classification system proprietors are all grouped with managers in a set of occupations labeled "owners, managers, and officials" (CPS:250–90). In contrast, the Canadian classification system categorizes proprietors on the basis of both the size and nature of their business. In the case of large and medium-sized firms, the owners are treated in a fashion very like that followed in the United States; they are categorized with managers. When the business is a small one, however, the owner is categorized with the main occupation associated with that business. For example, the owner of a small welding shop would not be classified as an owner-manager, but rather as a welder. From our own data and from available census tabulations, we have reason to believe that the proportion of respondents misclassified in the Canadian data (using the American data as a baseline) is small and is unlikely to seriously affect our findings.

THE VARIABLES

The variables which will be employed in the analysis are father's occupation (P), son's first-job occupation (S), time—based on the son's birth cohort—(T), and country of residence (C). The coding of father's and son's occupations has been discussed above. The cohorts employed in this analysis are five-year birth cohorts based on those who were born between 1908 and 1937. The American respondents were aged 25 to 54 at the time of the American study (1962), and the Canadian respondents were aged 36 to 65 at the time of the Canadian study. If we assume that age 16 is a rough approximation of the average age of labor-force entry for these respondents, then our analysis of origin to first-job mobility spans the years from 1924 to 1953.

The choice of son's first-job occupation for this type of analysis has been discussed by others (Hauser et al., 1975; McRoberts, 1975). Briefly, the purpose of doing this is to eliminate age (career) effects from consideration by taking the son's occupation at an identical point in the socioeconomic life cycle for each cohort, and hence focusing on period effects as the potential source of variability between cohorts. However, as Mason et al. (1973) have pointed out, three effects are confounded in cohort analysis: age (in our case career effects), period effects (changes due to historical events), and cohort effects (effects due to the specific composition of the cohort, e.g., the advantage a "small" cohort will have over a larger cohort in an expanding economy). We can, as noted above, remove the age effects from consideration by looking at first job. However, we cannot disentangle the effects of period and cohort.

THE MODELS FOR ANALYSIS

In the analysis of occupational mobility it is conventional to break the observed mobility into two analytically distinct components: structural mobility, and exchange or circulation mobility. The former refers to that mobility which occurs due to shifts from the fathers' generation to the sons' in the occupational structure as a result of shifts in the nature of occupational demand. The exchange or circulation mobility is seen as that mobility which would have occurred independently after the effects of intergenerational shifts in the marginal distributions have been taken into account or netted out. In a sense we view this as the "real" mobility—that is, the amount of mobility which would occur regardless of structural changes in the labor force in the intergenerational period. In addition to the intergenerational shifts in the occupational structure, when we look at a series of cohorts, it is also necessary to control for intercohort shifts in the occupational distributions of fathers and sons.

The models which we will be using for the intracountry trend-analysis are the same as those employed by Hauser et al. (1975) and are based on the log-linear models for contingency table analysis (see Goodman, 1971, 1972; and Bishop et al., 1975). The basic equation for these models can be expressed

$$\text{Log } F_{ijk} = u + u^p_i + u^s_j + u^t_k + u^{pt}_{ik} + u^{st}_{jk} + u^{ps}_{ij} + u^{pst}_{ijk} \quad (1)$$

Log F_{ijk} is the model-based maximum likelihood estimator of Log f_{ijk}, which is the logarithm of the observed cell value in the mobility table. If the expected values (Log F_{ijk}) fit the observed values (Log f_{ijk}) reasonably well, then we can say that the model fits the data (just as we would in regression analysis in finding a good fit between Y and $\hat{Y}$). The saturated model (Equation 1) will, of course, always provide an exact fit with the data. While it is not, as a result, an analytically interesting model, it is useful to look at the components of the equation in terms of their substantive meaning.

Log F_{ijk} represents the estimate of the set of observed values in the father-by-son-by-cohort matrix.

The terms u^p_i, u^s_j, and u^t_k reflect the effects on Log F_{ijk} of the distribution of father's occupation, son's occupation, and cohort size.

The terms u^{pt}_{ik}, u^{st}_{jk} represent the effects on Log F_{ijk} of the shift with respect to time (across cohorts) of father's occupational distribution and son's occupational distribution. Employed in combination, these control for the effects of differing degrees of structural mobility over time—a combination of the changes in the labor-force and cohort composition effects.

The term u^{ps}_{ij} represents the effect on Log F_{ijk} due to constant father-son interactions net of u^{pt}_{ik} and u^{st}_{jk}. In other words, this term represents the component of constant exchange mobility.

Finally, the term u^{pst}_{ijk} represents the effect on Log F_{ijk} of variability across cohorts due to changes in the patterns of exchange mobility.

Thus, the model in Equation 1 can be seen as decomposing the analysis of the changes over time in occupational mobility into the following components:

Observed Mobility (Log f_{ijk}) =
distributional effect ($u + u^p_i + u^s_j + u^t_k$)
+ the effects of change over time (across cohorts) in the distribution of father's occupation and son's occupation, and the effects of cohort composition ($u^{pt}_{ik} + u^{st}_{jk}$)
+ the effects of constant father-son interactions (u^{ps}_{ij})
+ the effects of intercohort variability in the father-son interactions (u^{pst}_{ijk}).

The model of substantive interest here is the model which omits the term u^{pst}_{ijk}:

$$\text{Log } F_{ijk} = u + u^p_i + u^s_j + u^t_k + u^{pt}_{ik} + u^{st}_{jk} + u^{ps}_{ij} \quad (2)$$

This model represents our attempt to fit the observed values (Log f_{ijk}) without the need for a term specifying intercohort variability in the father-son interactions. If the estimates under the model do in fact produce a reasonable fit with the observed values, we will conclude that there is no temporal variability in the nature of the father-son interactions. That is, we will conclude that a model of constant exchange mobility fits the data.

Finally, before we proceed to the data, a few words should be said on what is meant by fitting the data. The usual procedure with such models is to employ the likelihood ratio chi-square (G^2) as the measure of goodness of fit between the Log f_{ijk} and the Log F_{ijk} where

$$G^2 = 2 \Sigma f_{ijk} (\text{Log } f_{ijk} - \text{Log } F_{ijk})$$

This is calculated and referred to a χ^2 distribution with the appropriate degrees of freedom. However, it will be noted that the magnitude of G^2 is a function of two components: the degree of fit (Log f_{ijk} − Log F_{ijk}); and, holding the first term constant, f_{ijk}, which is a function of the sample size.

In addition, neither of these samples was drawn using a simple random design; in both cases the complex multistage stratified cluster designs have very significant design effects. As Fellegi (1978) has proven, design effects which are greater than unity will result in values of chi-square (and by extension G^2) which are substantial overestimates of the true value of chi-square under the null hypothesis. Or, to put it in another way, the estimates of p-values under the null will be substantially below the correct values. Fellegi (1978) shows that the estimates can be greatly improved by multiplying the chi-square estimates by the inverse of the design effect. We have done this here using an estimate of .57 as the inverse for the Canadian sample and .63 (see Hauser and

Featherman, 1977) for the American data. However, as Fellegi demonstrates, while this will very much reduce the underestimation of p-values, it does not entirely do away with it. His simulations show that the p-values will continue to be small with respect to their true values.

As a consequence of these two factors (very large samples giving great statistical power, and design effects on estimates), which are likely to result in substantively trivial effects being statistically significant in conventional terms, we do not intend to rely heavily on conventional tests in deciding whether or not there is a good fit of the model with the data.

In assessing fit we will look at a number of other criteria as well:

(1) The ratio of the G^2 for the model in question to the baseline G^2 gives the proportion of the total variability which can be accounted for by adding in the omitted terms in the baseline model. This ratio indicates the degree to which a more complex model would improve the fit over and above that already obtained.

(2) The index of dissimilarity Δ gives the percentage of cases which are misallocated by the model.

(3) Where we are comparing models with the same degrees of freedom across samples, we will employ, in addition to a comparison of the within-sample patterns on criteria one and two, a "normalized G^{2*}" ($G^{2*} = G^2 \times 1000/N$) (see Bishop et al., 1975:329–32 for details). Where the values on this normalized G^2 are very similar for identical models, we will conclude that the same model fits equally well in both samples.

(4) As a final check before deciding whether a model fits, we will examine the differences (residuals) between the observed and fitted values for our models in order to see if there is any pattern to them.

CANADIAN AND AMERICAN TRENDS IN MOBILITY

Table 3 presents in verbal form and in equations the models to be tested in this section. Table 4 presents the findings for the within-country tests of trends in Canada and the United States. In both Canada and the United States there are significant variations in the distribution of fathers' and sons' occupations (Models A1 and A2). There appears to be slightly less variability over time in the distribution of son's occupation than for father's in both countries.

Table 3. Equations and Hypotheses for Models in Table 4

Model A1

Saturated equation: Log $f_{ik} = u + u_i^P + u_k^T + u_{ik}^{PT}$

Father's occupation does not vary by cohort.

Ho: $u_{ik}^{PT} = 0$ for all (i,k)

Model A2

Saturated equation: Log $f_{jk} = u + u_j^S + u_k^T + u_{jk}^{ST}$

Son's occupation does not vary by cohort.

Ho: $u_{jk}^{ST} = 0$ for all (j,k)

Models B1, B2, C1, C2, D1

Saturated equation: Log $f_{ijk} = u + u_i^P + u_j^S + u_k^T + u_{ik}^{PT} + u_{jk}^{ST} + u_{ijk}^{PST}$

B1. Father's occupation varies by cohort, son's occupation varies by cohort, no father-son interaction.
Ho: $u_{ij}^{PS} = u_{ijk}^{PST} = 0$ for all (i,j,k)

B2. Father's occupation varies by cohort, son's occupation varies by cohort, constant father-son interaction.
Ho: $u_{ijk}^{PST} = 0$ for all (i,j,k)

C1. Father's occupation varies by cohort, son's occupation varies by cohort, no father-son interaction off of the main diagonal.
Ho: $u_{ij}^{PS} = u_{ijk}^{PST} = 0$ for all $i \neq j$ and all k

C2. Father's occupation varies by cohort, son's occupation varies by cohort, constant father-son interactions off of the main diagonal.
Ho: $u_{ijk}^{PST} = 0$ for all $i \neq j$ and all k

D1. No intercohort variation in father-son interactions in the main diagonal (constant inheritance).
Ho: $u_{ijk}^{PST} = 0$ for all $i = j$ and all k.

NOTE: P = Father's occupation, S = Son's first occupation, T = Cohort.

In Model B1, the baseline model, the only effects allowed (or controlled) for are the temporal changes in the father's and son's occupational distributions (variations which are exogenous with respect to our current concern). Not surprisingly, this model does not fit the data; indeed, it generates a very large G^2 value. In the subsequent models we will be concerned in part with the extent to which additional terms reduce this.

Model B2 is the constant father-son interaction model and represents the test of Hypothesis 1 for both countries. In essence, this model asserts that, net of shifts in the marginal distributions, the observed data can be adequately fit by assuming that there is no change over time in the

Table 4. Models Testing for Trends on Native-Born Males in Canada and Nonblack Males in the United States

		G^2		p		%[a]		Δ		G^{2*}	
Model	df	Can.	U.S.	Can.	U.S.	Can.	U.S.	Can.	U.S.	Can.	U.S.
A. Margins[b]											
1. P,T	25	57.8	111.7	.000	.000	—	—	5.3	4.4	17.8	14.6
2. S,T	25	99.2	157.5	.000	.000	—	—	5.8	5.1	30.7	20.6
B. All men											
1. PT,ST	150	1116.5	2949.0	.000	.000	100.0	100.0	21.6	23.5	344.7	384.7
2. PT,ST,PS	125	76.9	170.9	>.500	.003	6.9	5.8	5.2	5.0	23.8	22.3
C. Movers											
1. PT,ST	114	197.6	477.9	.000	.000	17.7	16.2	7.2	7.3	99.0	99.5
2. PT,ST,PS	95	57.8	122.3	>.500	.029	5.2	4.1	3.6	3.6	28.9	25.5
D. Nonmovers											
1. B2 vs. C2	30	19.1	48.6	>.500	.015	1.7	1.7	1.6	1.4	15.4	17.0

NOTE: P = Father's occupation, S = Son's first occupation, T = Cohort.
[a] % = (G^2model/G^2model B1) × 100.
[b] Models A1 and A2 are based on the P by T and S by T tables respectively.

pattern of father-son interactions. In the Canadian case, this model fits the data very well on all criteria: the G^2 is less than the associated degrees of freedom, the model accounts for all but 8.6% of the baseline G^2, and just over 5% of the cases are misallocated by the model. In the American case, however, the statistical fit is not as good, with a G^2 value which is, in conventional terms, large with respect to its degrees of freedom. Further examination of the data suggests that this is simply an artifact of the larger American sample. There is little difference between the two with respect to the normalized G^{2*}, the percentage of baseline chi-square unaccounted for, or in the percentage of cases misallocated by the model. Finally, when the tables of observed and fitted values were examined, we were unable to discern any trend in the distribution or magnitude of the deviations. In sum, the data very strongly indicate that there is no change with respect to time in the pattern of father-son interactions in either Canada or the United States.

Models C1 and C2 in Table 4 check the possibility that our models, dominated as they are by the interactions in the main diagonal, result in a masking of changes in the pattern of mobility by relative stability in inheritance. Models C1 and C2 remove from consideration the interactions in the main diagonal of the father-son matrix. Model C1 tests the hypothesis of quasi-perfect mobility, which simply asserts that, if one does not "inherit" one's father's occupation, then where one ends up is a random event. The model does not fit the data, indicating in both countries that where a son ends up in the occupational distribution is related to his father's occupation even when he does not remain in the same occupational category as his father.

Model C2 tests the hypothesis that the interactions in the mobility triangles which were established in Model C1 do not vary over time. This model fits the data very well in both countries (in Canada $p > .5$). Finally, Model D1, which by taking the difference between Model B2 (constant interactions) and Model C2 (constant mobility) focuses on the main diagonal, tests for constant inheritance. Again, the model fits the data very well for both countries.

Thus, the data suggest the following conclusions for both countries:

> No variation exists across cohorts with respect to the pattern of father-son interactions net of marginal effects (Model B2).
>
> No variation exists over time in the pattern of father to son occupational mobility (Model C2).
>
> No variation exists over time in the pattern of father to son occupational inheritance (Model D1).

In sum, the patterns of mobility and inheritance, net of structural changes, reveal no change over time. Thus, Hypothe-

sis 1 is supported in both Canada and the United States.

Is there any difference between the two countries with respect to the strength and nature of these patterns? The models presented in Table 5 represent a test of this hypothesis (Hypothesis 3). In this sequence of models the two P × S × T tables which were employed in the analysis in Table 4 have been combined into a single four-way table in which the variable C for country denotes Canada or the United States.[3] Our concern is whether or not we can fit a model which specifies no variability in father-son interactions with respect to time or country or both (that is: $u^{psc} = u^{pst} = u^{pstc} = 0$). If this model (A2, Table 5) can provide a reasonable fit with the observed data, we might well conclude that there is no difference between Canada and the United States with respect to circulation mobility.

Model A1 (Table 5) is the baseline model and controls for the variations in the distribution of fathers' and sons' occupations with respect to both time and country. We expect such variations (e.g., Table 2); however, they are in a sense exogenous to our present concern. The baseline variability (G^2) then consists of variability due to a series of components: constant father-son interactions; father-son interactions varying with time; father-son interactions varying with country; and father-son interactions which vary with both time and country. The remaining models in part A of Table 5 allow us to examine the proportion of the variability which is due to each of these sources and to test the degree to which any given component represents a statistically significant reduction in the baseline variability.

Model A2 includes the term for constant father-son interactions. The model fits the data very well, with a G^2 of only 295.3 with 275 degrees of freedom (p = .206) and an index of dissimilarity of only 5.6. Further, when we contrast Model A2 with Model A1 (Model A6) we can see that the addition of the PS term is both substantively important, accounting for over 90% of the baseline G^2, and statistically significant as well. Thus, a model of Canadian-American mobility which assumes no difference between the two countries with respect to patterns of exchange mobility will provide a good fit with the data. However, it is useful to test directly for differences due to time and/or country by adding these terms to our model and asking whether they provide any useful improvement in fit.

Model A3 (Table 5) includes the term for differences in the patterns of father-son interaction with respect to time. As we would expect, this term fails to produce a significant improvement in fit (Model A7). Model A5 adds in the term for intercountry differences in father-son patterns. As can be seen in Model A8, this term does produce what appears to be a statistically significant improvement in fit, accounting for a G^2 reduction of 49.1 for 25 degrees of freedom. However, while the improvement in fit is statistically significant, it is substantively trivial, accounting for only 1.2% of the baseline variability and improving the misallocation of cases by the model by less than 1%. Certainly, if these data show a difference between Canada and the United States in the patterns of association between father's and son's occupations, it is exceedingly small.

The models in parts B and C of Table 5 test, for intergenerational occupational movers and stayers, the same sequence of hypotheses as were tested in part A. For both movers and stayers the constant mobility/inheritance models (B2 and C2) provide very good fits with the data in both statistical and substantive terms. Further, the term for variations over time

[3] Readers may note that the models for Canada and the United States in Table 4 do not sum to the corresponding models in Table 5. For example, B1 Can + B1 U.S. from Table 4 does not sum to Model A1 in Table 5. This is due to the fact that the models in each table were estimated independently, and that the adjustments for design effects were applied independently to each model, resulting in some small inconsistencies from rounding. We used ECTA, Carleton University Version 2, to estimate our models. This is a version of the program written by Leo Goodman and Robert Fay, in the Department of Statistics at the University of Chicago. The program also incorporates modifications added at the University of Wisconsin and Carleton University.

Table 5. Models for Intercountry Differences for Native-Born Canadian Males and Nonblack American Males

Model	G^2	df	p	Δ	%[a]
A. All men					
1. PTC,STC	3983.9	300	.000	22.9	100.0
2. PTC,STC,PS	295.3	275	.206	5.6	7.4
3. PTC,STC,PST	143.1	150	>.500	3.5	3.6
4. PTC,STC,PSC	243.8	250	>.500	5.1	6.1
5. PTC,STC,PST,PSC	94.0	125	>.500	2.8	2.4
Contrasts					
6. A1 vs. A2	3688.5	25	.000	17.3	92.6
7. A2 vs. A3	152.3	125	.048	2.1	3.8
8. A3 vs. A5	49.1	25	.002	0.7	1.2
B. Movers					
1. PTC,STC	663.1	228	.000	7.2	16.6
2. PTC,STC,PS	227.3	209	.184	4.1	5.7
3. PTC,STC,PST	118.6	114	.375	2.8	3.0
4. PTC,STC,PSC	177.3	190	>.500	3.6	4.5
5. PTC,STC,PSC,PST	69.7	95	>.500	2.1	1.8
Contrasts					
6. B1 vs. B2	435.8	19	.000	3.2	10.9
7. B2 vs. B3	108.7	95	.164	1.3	2.7
8. B3 vs. B5	48.9	19	.000	0.7	1.2
C. Nonmovers (Ai vs. Bi)					
1. PTC,STC	3320.7	72	.000	15.6	83.4
2. PTC,STC,PS	68.0	66	.413	1.5	1.7
3. PTC,STC,PST	24.5	36	>.500	0.7	0.6
4. PTC,STC,PSC	66.5	60	.268	1.5	1.7
5. PTC,STC,PST,PSC	24.3	30	>.500	0.7	0.6
Contrasts					
6. C1 vs. C2	3252.7	6	.000	14.1	81.6
7. C2 vs. C3	43.6	30	.051	0.8	1.1
8. C3 vs. C5	0.2	6	>.500	0.0	0.0
D. Nonfarm ALL					
1. PTC,STC,PS	175.3	176	>.500	3.7	4.4
2. PTC,STC,PST	82.0	96	>.500	2.3	2.1
3. PTC,STC,PST,PSC	59.1	80	>.500	1.9	1.5
Contrasts					
5. D1 vs. D2	93.3	80	.147	1.4	2.3
6. D2 vs. D3	22.9	16	>.100	0.2	0.6
E. Farm outflow blocked					
1. PTC,STC,PS	211.5	220	>.500	4.1	5.3
2. PTC,STC,PST	100.3	120	>.500	2.5	2.5
3. PTC,STC,PST,PSC	70.3	100	>.500	2.1	1.8
Contrasts					
5. E1 vs. E2	111.2	100	.212	1.6	2.8
6. E2 vs. E3	30.0	20	>.070	0.4	0.7
F. Additional contrasts. No intercountry difference for:					
1. All Men (A8)	49.1	25	.002	0.7	1.2
2. All Farm (A8 vs. D6)	26.2	9	.000	0.5	0.6
3. Farm inflow and inheritance (E6 vs. D6)	7.1	4	.130	0.2	0.1
4. Farm outflow (A8 vs. E6)	19.1	5	.002	0.3	0.5

NOTE: P = Father's occupation, S = Son's first occupation, T = Cohort, C = Country.
[a] % = (G^2 model/G^2 model A1) × 100.

fails to add significantly to the fit in either case (see Models B7 and C7). However, the model testing for intercountry differences in intergenerational mobility (B8) does show that the addition of this term will produce a statistically significant improvement in the fit of the data, although as was the case earlier, the extent of the substantive improvement is questionable. Model C8, on the other hand, which tests

for intercountry differences in inheritance, is quite conclusive in showing that the addition of the term for variations over time contributes virtually nothing to fitting data.

At this point we can conclude with some certainty that Canada and the United States do not differ with respect to their patterns of intergenerational occupational inheritance. However, when we consider the question of intercountry differences in patterns of mobility, the evidence is mixed. On the one hand, we have results which suggest that the inclusion of a term for intercountry difference in our model will improve the fit significantly. On the other, we have the very small substantive improvement in the fit of the model achieved by adding the PSC term, and the lack of any discernable pattern of intercountry difference in the residuals of the models, all of which suggests that Canada and the United States do not differ in their patterns of mobility.[4]

However, in the Canadian case especially, farm origins constitute a very large proportion of the total. Indeed, for all but the youngest cohort of Canadians, over 30% had farm origins. Given the fact that farming was such a significant origin category, and given that the major rural-urban shift in Canada occurred during the latter part of the period at which we are looking (1925–1953), and at a time when the American rural-urban shift was tailing off, we conjectured that the difference might be due to the differences in farm-to-nonfarm occupational movements.[5] Indeed, if, as hypothesized by Lipset, Canada is a more closed society than the United States, the greater mobility flow in Canada engendered by the off-farm migration might well have the effect of masking the difference between the two countries.

In part D of Table 5, we present selected models in which all farm origins and destinations have been blocked. For those who have neither a farm occupation nor farm origins, the model of constant father-son patterns of association provides an excellent fit with the data, and additional terms provide no significant improvement in fit. From these models two conclusions may be drawn: when those with either farm origins or destinations are removed from consideration, there are, unambiguously, no differences between the Canadian and American patterns of father-son interactions; and to the extent that any differences do exist, they are located in either the farm-inflow or farm-outflow portions of the table.

In part E of Table 5, we present selected models in which the farm outflow cells (father's occupation, farm; son's occupation, nonfarm) have been blocked. There are, again, no substantively or statistically significant differences between the two countries. This clearly locates the source of statistical difference in the farm outflow cells of the matrices. This is directly tested in Model F4, which is the contrast of Model E6 and Model A8. The combination of these models allows us to partition the original G^2 due to intercountry differences of 49.1 with 25 degrees of freedom into a G^2 of 19.1 with 5 degrees of freedom due to differences in the farm outflow patterns between Canada and the United States (significant at $p = .002$), and a G^2 of 30.0 with 20 degrees of freedom due to differences between the two countries in all other patterns of father-son association, which is not statistically significant ($p = .070$). On purely statistical grounds, we would conclude that while similar in all other respects, the Canadian and American patterns of exchange mobility do differ with respect to the patterns of farm outflow. However, given the problems discussed earlier with respect to the potential overestimation of G^2, in combination with the facts that the addition of the PSC term improves the fit of the model to the data in a substantively trivial way and the examination of the residuals (even when attention is focused on the farm outflow cells) fails to reveal any pattern of difference, we conclude that the apparent differences are more likely due to "noise" than to anything substantive. We conclude that there is no difference

[4] Again, we would remind our readers that, due to the sampling design effects in our data, the observed values of G^2 are likely to be upwardly biased, with the consequence that the true value of the p's may be larger than those which are observed.

[5] We are grateful to Monica Boyd who suggested this possibility to us.

between the two countries in terms of either the patterns of mobility or inheritance.

CONCLUSIONS

In examining the patterns of interaction between father's occupation and son's first occupation over six five-year age cohorts in Canada and the United States, net of the effects of changes in the marginal distributions, we failed to find convincing evidence of a trend in these interactions or of differences in the patterns of interaction for the two countries.

This does not, of course, mean that there has been no change in the overall pattern of observed mobility in either country, nor does it mean that Canada and the United States have identical observed patterns of mobility. Rather, any changes or differences which do exist lie outside the basic pattern of father-to-son circulation mobility. We have already pointed out one potential source of observed differences, the rate of structural change in the labor force. This change is partially tapped in the marginal distributions of our mobility tables. However, these marginal shifts are confounded with cohort composition effects. Further, as others have noted (Duncan, 1966, 1968; Hauser et al., 1975), these distributions are not, due to factors such as fertility and mortality differing both by occupational group and by cohort (further cohort effects), representative, except in the crudest way, of real labor-force distributions. For these reasons we feel that a detailed analysis of these marginals is not likely to enlighten us further. Since the distributions are quite similar (especially when the nonfarm populations are considered), we believe that there are likely to be few differences between Canada and the United States in terms of the effects of structural mobility, although the possibility clearly deserves further exploration.

The literature comparing Canada and the United States led us to expect that some differences would exist. Our findings concerning the nature of circulation mobility in the two countries speak directly to Naegele (1961) and Lipset's (1963) functionalist discussions of values and value orientations. Taking as given a value difference between Canada and the United States, our analysis suggests two possibilities: first, value differences were in the past and are today not of sufficient magnitude to affect mobility; or, second, value differences (and hence values themselves) are spurious with respect to mobility—i.e., value differences, even if they can be shown to exist, have no effect on circulation mobility.

In the case of Porter's (1965) argument that the failure of Canadian educational institutions to become more open to opportunity has lowered the rates of circulation mobility, implications from our analyses are both different and rather more interesting. The factual basis of Porter's argument is sound: not only are Canadians less well educated than Americans, but evidence has since been produced to show that their educational aspirations, at least, are more strongly tied to social origins than is the case for Americans (Gilbert, 1973:71–125). The fact that these differences do not have any effects on exchange mobility may be due to a number of factors. The most likely of these is that the way in which educational attainments are translated into occupational positions is different in the two countries. Due to limitations in the data, it is not possible to examine this possibility at this time. However, we feel that it is an important hypothesis for future work on Canadian-American differences.

Finally, while we have found little evidence of either trend or difference in comparing Canada and the United States with respect to circulation mobility, it should be remembered that this analysis is based on limited samples of the populations of the two countries. In Canada we have looked only at native-born males, and in the United States only at nonblack males. Groups such as women, migrants, and blacks have all been excluded from consideration. If these groups were considered, the findings might be quite different. Further, in this analysis we have only been concerned with the broad patterns of father-son interaction: mobility and inheritance. It is possible that using the methods more recently developed by Hauser, one would find some trends and

some intercountry differences within very specific occupational groups.

APPENDIX A

OCCUPATIONAL CODING

The first step in our procedure was to attempt to match each of the 1960 Census Occupational Titles (COT) with an equivalent title from the Canadian Classification and Dictionary of Occupations (CCDO). This in turn would allow us to assign each COT-CCDO match to a category in a 16-fold PPM classification of the CCDO titles themselves (see Pineo et al., 1977, esp. Appendix).

The census titles which presented problems in the initial mapping were of three types. First, ten COT titles with CCDO matches, when assigned to the PPM groups appropriate for their respective CCDO titles, appeared to be misallocated when their COT groups were compared to the PPM groups. For example, COT code 164, radio operators, was matched with CCDO code 9551, radio and TV broadcasting equipment operators. However, CCDO code 9551 falls in PPM group 10, skilled crafts and trades, while COT code 164 falls in the COT major group professionals, technicians, and kindred workers, which in turn seems more closely matched by PPM group 5, technicians, in this instance.

Secondly, some COT titles did not have clear CCDO matches. For example, the COT codes 184, teachers, n.e.c., and 192, technicians, n.e.c., do not have equivalents.

The third type of problem involved COT titles that encompassed two or more CCDO titles. Here, for example, the COT code 105, lawyers and judges, was matched by the CCDO codes 2343, lawyers, and 2342, judges.

Resolution of these problems followed one of three lines. In the case of apparently misallocated COT titles in the PPM scale, no changes were made except to shift two COT titles into the farmers and farm workers category and to shift one COT title from the farm group into the unskilled laborer category. Where we had COT titles without apparent CCDO matches, all of the problem cases consisted of COT titles qualified by the n.e.c. (not elsewhere classified) code. These residual categories, while not specifically matched by CCDO titles, were not problematic, as they could, in all but one case, be uniquely assigned directly to one of the PPM groups (i.e., foremen n.e.c. could be assigned directly to PPM group 8; foremen). The single difficulty was the COT code 195, professional technical and kindred workers n.e.c. This group, which straddles several of the PPM categories (groups 1, 2, 4, and 5), was judged, after an examination of the more detailed titles subsumed under it, to be best assigned to groups 4 and 5, semiprofessionals and technicians.

The third type of problem in the initial mapping—the COT titles that clearly encompassed two or more CCDO titles—was dealt with in one of three ways. First, where the several CCDO titles subsumed under a single COT title fell into a single PPM category, the COT title was simply assigned to that category without a unique match in the CCDO listing having been made. The second method of dealing with multiple CCDO matches was to collapse the 16-fold Pineo, Porter, and McRoberts (1977) scale. In fact, early in our research we became aware that the likelihood of readily matching the American COT classification to the PPM 16-category classification was very small, given the level of detail in the 16-fold scale. As a result we anticipated the necessity of collapsing the scale to facilitate the construction of comparable national occupational categories. The initial collapsing, based upon both the necessities of comparable cross-national scales and substantive importance, reduced the occupational classification to eight categories: professionals, semiprofessionals, and technicians; managers and administrators; upper white-collar workers; upper blue-collar workers; lower white-collar workers; lower blue-collar workers; and, finally, farmers and farm laborers.

This collapsing eliminated a large number of the problems in matching the COT titles to CCDO titles and then to PPM groups, as many of the COT titles which prior to the recode had straddled two or more PPM groups now fell within one of the eight new categories. For example, the COT code 105, lawyers and judges, was matched by the CCDO code 2343, lawyers, which fell into PPM group 1, self-employed professionals, and judges, CCDO code 2341, which fell into PPM group 2, employed professionals. Once the PPM groups 1 and 2 were collapsed, the COT code 105 could be assigned to the collapsed group without misallocation.

At this point, we had assigned each COT title to a unique category of the 8-fold PPM-based scale. The first check made was to generate and compare frequency counts for the three occupational variables in each data set: father's occupation when respondent was 16, respondent's first job, and respondent's present job. In comparing the two countries on comparable samples (males, 25–64, in the labor force) our criteria of comparison were the assumptions that the occupational structures of the two countries would not differ greatly and that any large or systematic differences that did appear should be treated as a potential coding problem until proven otherwise.

The first comparison of the two data sets on each of the three variables showed several very large differences in the distribution of the respective labor forces across the eight occupational categories. The pattern of the differences, however, appeared to be fairly systematic: in 5 of the 8 categories of respondent's present job, the two countries differed by 2.5% or less. In the three problematic categories the differences were quite large. In the category for high and middle managers, the United States almost doubled the proportion of Canada (15.98 vs. 8.89). The two other groups that were not close in their relative frequencies were the categories for upper white-collar and lower white-collar workers.

Our first response to these apparent problems in creating comparable occupational classifications was to recheck our initial matching procedure. This was done by taking each CCDO code in the three problem categories and matching it in the detailed 1970 *Alphabetical Index of Industries and Occupations* and in turn determining the appropriate census code for each CCDO code. It was felt that this reversal of our initial mapping (from mapping the American onto

the Canadian to mapping the Canadian on the American) would identify the persistent problems in our classification. In the end, this cross-check resulted in only a few minor changes in the original assignments for the census codes; and a check on the frequency distributions for respondent's present job showed little appreciable change in the three problem areas. On the basis of this result, we concluded that the source of the problem we were experiencing lay not in our matching procedure, but in more fundamental differences in the two basic occupational classifications, the COT and CCDO.

Following this direction, we were able to determine that the great difference initially found between the two countries in the proportion of their labor forces allocated to the managers' and proprietors category of the 8-fold classification was due to differences in the treatment of proprietors in the respective classifications. In the COT classification, proprietor or owner is recognized as a unique occupation and is generally grouped with managerial occupations in being assigned a code. Thus the majority of proprietors fall in the group of COT codes ranging from 250 to 290, which are also the managers and official codes. In the allocation of the COT titles to the PPM scale, the majority of COT codes which subsume proprietors were translated into CCDO managerial codes, and thus into the group for high and middle managers.

In contrast to the American procedure, the CCDO does not recognize proprietor as an occupation, but rather classifies a proprietor by the nature of his or her work and the size of the business owned. Thus a barber shop owner is coded as a barber, and only the owners of rather large firms are classified with managers in the CCDO. This has the effect of spreading proprietors in the Canadian classification over virtually all of the occupational levels, from semiskilled blue collar to white collar to managerial and professional.

The overall effect of the differing treatment of proprietors (including self-employed with no employees as well as self-employed with employees) is that they appear as a large group in the managerial category of the 8-fold PPM scale when we apply it to American data, while in contrast they "appear" as a number of small groups spread across several of the PPM categories when applied to Canadian data.

The differences due to the differing treatment of proprietors were a problem that we were unable to resolve through recoding procedures in the COT-CCDO matching. As a result we chose to collapse further the 8-fold scale in order to minimize the possible effects of this problem. In doing so we combined the professionals and semiprofessionals into a single group and combined the managerial and upper-white-collar groups into another single group (managers and proprietors), designating the group formerly titled "lower white collar" as simply "white collar."

The six-fold classification upon which we finally decided is presented in Table 2, broken down by country and by age cohorts, for respondent's first job. While some noticeable differences between the countries do exist, we feel that these are either consistent with the expected differences, based upon other knowledge of the two countries, or, where the differences are due to coding error, that they are not large enough to be of great significance. In short, we feel that, as a first approximation, the matching of the two occupational classifications is adequate as a base for our analysis.

REFERENCES

Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul W. Holland
1975 Discrete Multivariate Analysis: Theory and Practice. Cambridge: The MIT Press.

Blau, Peter M. and Otis Dudley Duncan
1967 The American Occupational Structure. New York: Wiley.

Blishen, Bernard R. and Hugh A. McRoberts
1976 "A revised socioeconomic index for occupations in Canada." Canadian Review of Sociology and Anthropology 13(1):71–9.

Boyd, Monica and Hugh A. McRoberts
1974 "Design of the 1973 Canadian national mobility study on occupational and educational changes in a generation." Paper presented to the North Central Sociological Association Meeting, Windsor, Canada, May.

Boyd, Monica and John Porter
1977 "Immigration and ethnic effects on occupational status." Presented to the International Sociological Association Research Committee on Stratification Conference, Dublin, Ireland.

Broom, Leonard and F. Lancaster Jones
1969a "Career mobility in three societies." American Sociological Review 34:650–8.
1969b "Father-to-son mobility: Australia in comparative perspective." American Journal of Sociology 74:333–42.

Canada
1971 Occupational Classification Manual, Census of Canada, 1971. Based on Canadian Classification and Dictionary of Occupations. Volumes I and II. Ottawa: Information Canada.

Dofny, Jacques and Muriel Garon-Audy
1967 "Mobilites professionnelles au Quebec." Sociologie et Societes 1:277–301.

Duncan, Otis Dudley
1966 "Methodological issues in the analysis of social mobility." Pp. 51–77 in Neil J. Smelser and Seymour Martin Lipset (eds.), Social Structure and Mobility in Economic Development. Chicago: Aldine.
1968 "Social stratification and mobility: problems in the measurement of a trend." Pp. 675–719 in Eleanor B. Sheldon and Wilbert E. Moore (eds.), Indicators of Social Change. New York: Russel Sage Foundation.

Featherman, David L. and Robert M. Hauser
1978 Opportunity and Change. New York: Academic Press.

Fellegi, I. P.
1978 "Approximate tests of independence and goodness of fit based on stratified and

APPENDIX B

Procedure for Generating the Six-Fold Occupational Classification from Census Occupational Codes

GROUP 1 Professionals and Semiprofessionals														
	000	010	012	013	014	015	020	021	022	023	030	031		
	032	034	035	040	041	042	043	045	050	051	052	053		
	054	060	070	071	072	073	074	075	080	081	082	083		
	084	085	090	091	093	102	103	105	111	120	130	131		
	134	135	140	145	150	152	153	160	161	162	163	165		
	170	171	172	173	174	175	181	182	183	184	185	190		
	191	192	193	194	195	253	265	303	420	642				
GROUP 2 Managers and Proprietors														
	092	101	104	151	154	180	222	250	251	254	260	262		
	270	275	280	285	290	301	310	321	333	342	345	352		
	380	385	393	395	450	493	821	832	852	853	854			
GROUP 3 White Collar														
	302	304	305	312	313	315	320	323	324	325	340			
	341	343	350	351	353	354	360	370	381	382	383			
	390	394	650	801	802	804	810	812	814	815	830			
	840	842	843	874	890									
GROUP 4 Upper Blue Collar														
	164	252	314	401	402	403	404	405	410	411	413	414	415	421
	423	424	425	430	431	432	434	435	444	451	452	453	454	460
	461	465	470	471	472	473	474	475	480	490	491	492	494	495
	501	502	503	505	510	512	513	514	515	520	521	523	524	525
	530	535	545	601	602	603	604	605	610	612	613	614	615	620
	621	630	631	634	640	641	643	645	651	652	653	670		
	671	672	673	674	675	680	690	691	692	694	695	701		
	704	705	712	713	720	721	775	825	831	850	962	970		
GROUP 5 Lower Blue Collar														
	504	632	635	654	685	693	703	710	714	715	803			
	813	820	823	824	834	835	841	851	860	960	963			
	964	965	971	972	973	985								
GROUP 6 Farmers and Farm Laborers														
	200	901	902	903	905									

multi-stage samples." Survey Methodology 4(2):29–56.

Gilbert, Sid N.
1973 Educational and Occupational Aspirations of Ontario High School Students: A Multivariate Analysis. Doctoral Dissertation, Carleton University, Sociology, Ottawa, Canada.

Goldthorpe, John H., Catrione Llewellyn, and Clive Payne
1980 Social Mobility and Social Class in Modern Britain. Oxford: Clarendon Press.

Goodman, Leo A.
1971 "The analysis of multidimensional contingency tables: stepwise procedures and direct estimation methods for building models for multiple classifications." Technometrics 13:33–61.
1972 "Some multiplicative models for the analysis of cross-classified data." Pp. 649–96 in Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability. Berkeley: University of California Press.

Hauser, Robert M. and David L. Featherman
1977 The Process of Stratification. New York: Academic Press.

Hauser, Robert M., John N. Koffel, Harry P. Travis, and Peter J. Dickinson
1975 "Temporal change in occupational mobility: evidence for men in the United States." American Sociological Review 40:279–97.

Hope, Keith
1974 "Trends in the openness of British society in the present century." Paper presented at the Mathematical Social Sciences Board Conference on Measurement and Models in Comparative Social Stratification, Toronto, Canada, August.

Lipset, Seymour Martin
1963 The First New Nation. Garden City: Anchor Books.

Lipset, Seymour Martin and Reinhard Bendix
1959 Social Mobility in Industrial Society. Berkeley: University of California Press.

Mason, Karen Oppenheim, William Mason, H. H. Winsborough, and W. Kenneth Poole
1973 "Some methodological issues in cohort analysis of archival data." American Sociological Review 38:292–58.

McRoberts, Hugh A.
1975 Social Stratification in Canada: A Preliminary Analysis. Doctoral Dissertation,

Carleton University, Sociology, Ottawa, Canada.

Miller, S. M.
1960 "Comparative social mobility." Current Sociology 9:1–89.

Naegale, Kaspar D.
1961 "Canadian society: some reflections." Pp. 1–53 in Bernard R. Blishen, Frank E. Jones, Kaspar D. Naegele, and John Porter (eds.), Canadian Society. Toronto: Macmillan of Canada.

Pineo, Peter C., John Porter, and Hugh A. McRoberts
1977 "The 1971 Census and the socioeconomic classification of occupations." Canadian Review of Sociology and Anthropology 14:91–102.

Porter, John
1965 The Vertical Mosaic. Toronto: University of Toronto Press.

Treiman, Donald J. and Kermit Terrell
1975 "The process of status attainment in the United States and Great Britain." American Journal of Sociology 81:563–83.

Stone, Leroy D.
1967 Urban Development in Canada. Ottawa: Queen's Printer.

U.S. Bureau of the Census
1960 Classified Index of Occupations. Washington: U.S. Government Printing Office.

ORGANIZATIONAL CONTEXT AND SCIENTIFIC PRODUCTIVITY*

J. SCOTT LONG
Washington State University
ROBERT MCGINNIS
Cornell University

An earlier study found that while a scientist's productivity does not affect the prestige of the academic position obtained, the prestige of the position does affect later productivity. In this paper consideration of contextual effects is broadened to include differing organizational contexts of scientific employment. Chances of obtaining employment in a particular context are not strongly affected by productivity. Once employment is obtained in a specific context, individual levels of productivity soon conform to characteristics of that context. These results do not support the idea that scientists are allocated to organizational contexts on the basis of their scientific contributions. Past research indicating that the most productive scientists are recruited to academic locations may have confused the cause of a scientist working in a given context with the effect of working in that context.

Studies of the scientific career are generally limited to scientists who hold faculty positions in Ph.D.-granting institutions, implicitly suggesting that scientists outside academia are not scientists in the same sense as those in academia. For example, Marcson (1960) and Kornhauser (1962) contrast professional scientists with organizational scientists; Hagstrom (1965) contrasts the "scientific community" with industrial research settings. The implicit assumption is that real science is performed only in the university. This assumption is based at least in part on a second assumption: the normative structure of science—as proposed by Merton (1973) and others (Hagstrom, 1965; Storer, 1966)—operates differentially among sectors of employment.

According to the normative structure of science, contributions to the body of scientific knowledge are the fundamental criteria of evaluation. The scientist is viewed as operating in a process of exchange—knowledge for prestige and recognition (Hagstrom, 1965:13). The norm of communality requires that knowledge be made public for the gain of the community and not directly for that of the individual or employer. Such a model of science is based on science as performed in an academic setting (cf. Sklair, 1972), not in an industrial, developmental laboratory.

In nonacademic locations, particularly industry, the criteria are thought to be vastly different. Kornhauser (1962:13) claims that the requirements for success in an organization commonly clash with those for success in science as a profession. Industrial locations may control publications (although there are important exceptions), if not stop them altogether. The emphasis on the free exchange of information that is suggested by the functional theory of science is replaced by the more pragmatic logic of business. Publications must be delayed until their results are adequately covered by patents (cf. Cotgrove and Box, 1970:19).

Most studies of nonacademic scientists (Marcson, 1960; Kornhauser, 1962; Glaser, 1964) have focused on the difficulties that scientists may have in adapting to the perhaps deviant environment of industry: difficulties that arise from a conflict between the norms which scientists acquired during their graduate edu-

* Direct all correspondence to J. Scott Long, Department of Sociology, Washington State University, Pullman, WA 99164. We would like to thank Lowell Hargens, Barabara Reskin and two reviewers for their comments on an earlier version of this paper. S. Gunason provided valuable editorial advice. Support was provided by NIH Contract NO1-OD-2175 awarded to Cornell's Program on Social Analysis of Science Systems.

cation and the expectations of their industrial employers. Cotgrove and Box (1970) indicate that this conflict is not necessarily as strong as previously argued, and may be almost nonexistent due to the differing motives of those who work in industry as opposed to academia. The suggestion of differing motives between nonacademic and academic scientists has also been noted by Krohn (1971:107), who states, "The argument has been that universities, government and industry respectively shape their research and people who do it. This socialization begins with the process of selective recruitment, whereby institutions and people reciprocally influence each other."

Regardless of the extent of conflict between nonacademic scientists and their employers, the choice of the organizational context of employment is an extremely important one. The context determines the research setting and facilities. It affects not only professional prestige, since academia is generally considered the most prestigious and desirable location and industry the least (Zuckerman, 1970), but also the opportunities for research in terms of freedom to publish and intellectual autonomy. Further, the choice of a nonacademic career is often final (Harmon, 1968:10); industrial employment almost assures that a return to academia will be blocked (Caplow and McGee, 1958:150).

Organizational context may be related to scientific productivity in a variety of ways. First, as Reskin (1977:493) has noted, the different goals of various locations may either hinder or promote productivity. For example, promotion in research universities is generally dependent upon publications. In industrial settings, publications are generally not criteria for promotion. The fact that such locations do not encourage publications, but rather tolerate them, may result in an increasingly negative effect of nonacademic locations on levels of publication through a process of reinforcement. Second, several studies (Caplow and McGee, 1958:150; Cotgrove and Box, 1970:2; Klaw, 1968:78; Marcson, 1960:54) have suggested that nonacademic locations have difficulties in recruiting the best graduates and must settle for the second echelon. If so, the lesser abilities of nonacademic scientists would contribute to their lower productivity.

While numerous studies have contributed to our understanding of stratification within the academic sector (Allison, 1976; Cole and Cole, 1973; Gaston, 1978; Hagstrom, 1971; Hargens and Hagstrom, 1981; Long, 1978; Long et al., 1979), very few have emerged to explicate stratification in science as a whole, including both the academic and nonacademic sectors (Reskin, 1976, 1977, 1979). Yet to understand the entire social system of science, the significant and growing nonacademic sector must be considered. Further, to fully understand stratification within academia, the process leading to an academic career as opposed to a career in industry or government must be understood. To fill a gap in past research we provide a longitudinal analysis of productivity and position for a representative sample of both academic and nonacademic scientists. Three major questions are addressed. First, what determines the allocation of scientists to various organizational contexts? Specifically, do more productive scientists obtain positions in more prestigious locations? Second, to what extent does organizational context influence a scientist's productivity? Does a scientist begin to conform to the perhaps alien expectations of a nonacademic environment? Finally, are differences in productivity primarily the consequence of recruiting the more productive scientists in some sectors, or are they the result of working in different organizational contexts?

DATA AND MEASUREMENT

Analysis is based on the population of male biochemists who obtained their doctorates in the fiscal years 1957, 1958, 1962, and 1963. Female doctorates were excluded from consideration due to their small number and the difficulty encountered in obtaining complete biographic information on their careers. Biographic information was coded from the 10th through 13th editions of *American Men (and Women) of Science*. Complete in-

formation on educational and occupational experiences was obtained for 557 (83%) of the 668 males who obtained their degrees during this period. Prestige of the doctoral department was measured with the 3-digit ratings of faculty quality of biochemistry departments, a partial listing of which appeared in Cartter (1966). The prestige of postdoctoral appointments in graduate departments was based on a weighted average of the Roose and Andersen (1970) prestige scores for bioscience departments in the fellowship institution.[1] An average was used since the specific department of the postdoctoral appointment was not always known. The prestige scores range from a low of 1.00 to a high of 5.00. Whether a graduate department was located in a college of agriculture was also coded, a distinction in keeping with biochemistry's twin origins in medical and agricultural science (cf. Bernal, [1954] 1971:886ff.).

While the prestige of the doctoral department is often considered in studies of the scientific career, characteristics of a student's dissertation supervisor, hereafter referred to as the mentor, are generally ignored. Yet the mentor may be an important influence both in training and sponsoring the student. Using the American Chemical Society's *Directory of Graduate Research, Dissertation Abstracts*, and inquiries to specific universities, the name of the mentor was obtained for all but two of our sample. Following the work of Reskin (1979), two basic characteristics of the mentor were isolated: eminence and performance.

Performance reflects current (at the time of the student's doctorate) research productivity. While several measures of performance were collected, the indicator used is the number of citations (to both junior and senior authored papers) received by the mentor to papers published in the three-year period beginning with the student's doctorate. Citations were coded from the 1961 and 1966 volumes of *Science Citation Index*. Scientific eminence, while stemming from scientific performance, may lag performance and persist after performance declines (Zuckerman and Merton, 1971:81 and 1972:325; as cited by Reskin, 1979:131). A number of indicators of eminence were coded, ranging from election to the National Academy of Science to the numbers of honorary degrees, postdoctoral fellowships, or scientific awards received. The measure used in the following analysis is a weighted count of prestigious awards such as the Nobel Prize and election to the National Academy of Science (weighted 4), honorary degrees (weighted 3), prestigious fellowships such as a Fullbright or Guggenheim (weighted 2), and less prestigious awards such as the Borden or Vaughan Awards (weighted 1). While more refined weightings based on prestige of the awards, such as those presented by Cole and Cole (1973:270–5) for physics, would be preferable, such a list is not available for biochemistry; and there were too many differences in the specific awards in biochemistry compared to physics to usefully apply the Coles' list. Details on the construction of this measure are presented in Long and McGinnis (1981).

Two measures of the employment experiences of the mentor were also coded. The contexts in which the mentor has worked may influence the context in which the student obtains employment. These measures are described below.

Productivity of the sample members was measured using counts of both publications and citations to them. *Chemical Abstracts* (1955–1973) was used to locate the articles published by the sample members, whether or not they were the senior authors. Citations to these articles were coded from *Science Citation Index* (Vols. 1961, 1964, 1966, 1968, 1970, 1972, and 1974). On multiple-authored papers where the cohort member was not the first author, the name of the first author was used to locate citations to junior authored papers (cf. Long et al., 1980 for details on this problem); thus downward bias in counts for scientists who were predominantly junior authors was avoided. For a given year in the scientist's career, the publication measure reflects publications in a three-year period ending in that year. The citation measure for that year is restricted to citations to papers published in that three-year period. Since coverage of *Science Citation Index* and *Chemical Ab-*

---

[1] The complete Roose and Andersen prestige scores were kindly provided by Charles J. Andersen.

stracts increased during the period covered by our analyses, counts were standardized within years of the Ph.D. For further details, see Long (1978).

CLASSIFICATION OF PRIMARY WORK ACTIVITY AND ORGANIZATIONAL SECTOR

A clear classification of organizational contexts is necessary to study the process by which scientists are allocated to various contexts and the effects of these contexts on scientific productivity. Two basic dimensions of context must be considered: employment sector and primary work activity.

The first dimension of organizational context is sector of employment. In his review of research settings, Marcson (1972) identifies the four sectors of government, industry, nonprofit, and university, with a secondary distinction between applied and basic research. Pelz and Andrews (1966) compare Ph.D.s in development-oriented laboratories (mainly industry and parts of government) with those in research-oriented laboratories (academia and parts of government). Harmon (1968) uses four divisions: academia, business and industry, government, and all other locations. Reskin (1977) classifies organizations in terms of their research orientations with two dummy variables: the first distinguishing university (but not college) locations from other locations; the second distinguishing research-oriented contexts from other contexts. Research-oriented contexts include universities, government laboratories, and nonprofit institutes. Unlike the other studies, Reskin notes the fundamental importance of the distinction between research universities and four-year colleges.

The second dimension concerns the work activities or roles which scientists perform. The three major activities distinguished by Harmon (1965) and later analyzed by Zuckerman and Merton (1972) are teaching, research, and administration. Most scientists spend some time in each of these roles, and accordingly one would prefer to know the percentages of time a scientist devotes to each of these activities. Unfortunately, such information is not accessible from secondary sources such as those used in this study. Accordingly, classifications into what appeared to be modal work activities were made on the basis of the job titles given in *American Men and Women of Science*. Within industrial settings and research laboratories titles such as lab director, section head, and vice president for research were used to indicate administrative duties. Titles such as associate scientist, research scientist and biochemist were taken to indicate primarily research activities. Faculty positions, which often combine both teaching and research roles, were classified on the basis of job titles such as assistant, associate, or full professor.

The distinctions among research, teaching, and administration are not in themselves completely adequate. Clearly, teaching in a research university has a strong and perhaps dominant research component, while teaching in a four-year liberal arts college may have a negligible research component. Research carried out in a research university will be quite different from that done in most industrial laboratories. These distinctions should reflect the extent to which various organizational settings encourage the publication of research results. Within the academic sector, universities strongly encourage publication, if not outright demand it, while in four-year colleges and nonresearch universities it may be simply tolerated. Among research laboratories, locations such as academic research centers, private-nonprofit laboratories, and those exceptional industrial laboratories (e.g., Bell Laboratories and Sandia Laboratory of AT&T) encourage publication, while the majority of industrial laboratories and some federal laboratories restrict publication. Given the variety of research settings within some of the largest employers (e.g., General Electric), it is impossible to be certain if a scientist is in a location which encourages or discourages publication. Yet, by following the suggestions of past research, our division of employers into those who encourage publication and those who discourage publication should be quite accurate.

In our final classification, both primary work activity and sector are considered. The resulting classes are referred to as organizational contexts. Six contexts are

distinguished, which are listed below along with abbreviations in parentheses (see Table 1).

(1) Faculty in a research university (FACPUB), where research universities are defined as those universities offering doctoral degrees in bioscience. Nearly all of these were rated in the Roose and Andersen (1970) study.
(2) Faculty in nonresearch universities and four-year colleges, where publication is not likely to be a job requirement (FACNPB). These are primarily four-year colleges, and accordingly this context will sometimes be referred to as teaching in a college.
(3) Research in locations where publication is assumed to be encouraged (RESPUB). These consist primarily of academic, private nonprofit, and governmental laboratories. In addition, this context includes those outstanding industrial laboratories in which publication is encouraged. This category does not include either prestigious academic research appointments or short-term academic research positions held prior to a tenure-track faculty appointment. Job titles include senior scientist, research associate, biochemist, scientist, and similar designations.
(4) Research in locations where publication is not assumed to be encouraged (RESNPB), primarily industrial laboratories. For convenience this category will sometimes be referred to as industrial research.
(5) Administration in any sector (ADMIN). During the period of the career that our analyses will focus upon; this is primarily the administration of research in industrial laboratories.
(6) Since the postdoctoral fellowship is an important career step for many biochemists (see below for futher details), the category of postdoctoral fellow is included (FEL).

Several fundamental characteristics of the careers of biochemists are illustrated by the data in Table 1. First, the postdoctoral fellowship is the modal start of the career. Second, upon completion of the fellowship, a majority of fellows end up in faculty positions at research universities, followed by research positions in locations which encourage publication. Third, by the ninth year of the career there has been movement out of both types of research locations and an increase in the category of administration. In analyses not presented here, it is found that very little change among organizational contexts occurs, beyond what is necessary for the observed shift in the marginals. What change does occur is largely the increase in administration due to those who move from research to administrative roles.

THE ALLOCATION OF POSTDOCTORAL FELLOWSHIPS

Given the importance of the postdoctoral fellowship in terms of its numbers as noted above, its function in the training of biomedical scientists (cf. Curtis, 1969;

Table 1. Percentages of Scientists in Various Organizational Contexts over Time

Code	Category	% Immediately after Ph.D.	% 1st Postdoctoral, Postfellowship Location	% at Ninth Year of Career
FACPUB	Faculty in a research university	18.9	51.9	51.7
FACNPB	Teaching in a nonresearch university, primarily 4-year colleges	4.0	6.4	8.5
RESPUB	Research in locations where publication is encouraged	13.2	21.8	17.5
	% of category in academia	(34.7)	(43.7)	(33.7)
	% of category in government	(38.9)	(34.9)	(38.9)
RESNPB	Research in locations where publication is not encouraged (most industry and some government locations)	10.3	13.9	11.6
	% of category in industry	(100.0)	(97.4)	(92.1)
ADMIN	Administration in any location	4.4	5.9	10.3
FEL	Postdoctoral fellowships in any location, primarily research universities	49.2	0.0	0.4
		N = 545	N = 545	N = 543[a]

[a] Ns differ due to varying numbers of scientists excluded because of atypical work activities.

NAS, 1974), and its effects on later scientific achievement, the receipt of a postdoctoral fellowship is an important first step in many scientific careers. Research by McGinnis et al. (1981) has examined the allocation of fellowships along with the consequences of receiving a fellowship. The most significant result of their analyses is the lack of effect of predoctoral productivity on the odds of receiving a fellowship, whether productivity is measured by publications or citations. Thus, at least in biochemistry where postdoctoral fellowships are common, fellowships cannot be viewed as rewards for the demonstrated ability (to the extent that ability is reflected by published work) of young Ph.D.s. Rather, a combination of structural and personal effects dominate. Students from agricultural schools are much less likely to obtain fellowships. Young and unmarried students are more likely to obtain fellowships than older and married graduates. Finally, students from prestigious departments are more likely to receive a fellowship than those from less prestigious departments, independently of characteristics of the mentor.

ENTRANCE INTO THE FIRST JOB

The organizational context of a scientist's first job is of critical importance to his entire career. It provides both the physical resources with which to conduct research and the social context within which to work. Given the barriers to mobility which exist between the academic and nonacademic sectors (cf. Harmon, 1968), the first job is also important for its permanence.

To determine the extent to which various characteristics of a scientist determine the organizational context of his first job, the effects of three sets of variables are assessed. First, characteristics of the educational experience are considered, including the prestige of the doctoral department, the eminence and performance of the mentor, and the type of work experience had by the mentor. Second, the productivity of the student at the time he obtained the first position is considered, measured both with numbers of publications and citations. Finally, since the fellowship plays such an important role in biochemistry, the fellowship itself is examined as a factor determining the choice among organizational contexts. This is done both by considering the simple fact of having a fellowship and by taking into account changes in productivity and prestige which result from the fellowship. Thus, if the fellowship was held in a graduate department, the prestige of that department replaces the prestige of the doctoral department in explaining the allocation of positions. And for those with fellowships, the level of productivity at the end of the fellowship replaces productivity at the time the doctorate was completed.

Our problem is to determine the effects of both continuous and discrete independent variables on a nominal dependent variable with five categories. This problem is ideally suited to multinomial logit analysis (cf. Domencich and McFadden, 1975; Nerlove and Press, 1973). Since sociologists are generally unfamiliar with this technique, a brief overview may be useful. Simply put, multinomial logit analysis estimates the effects of various independent variables on the odds of one outcome compared to another outcome. It is a generalization of log-linear models for dependent variables presented by Goodman (1971, 1972), allowing for continuous independent variables. When there are two outcomes, there is only one possible comparison. When there are more outcomes, such as in the current problem, a series of comparisions can be made. Thus, FACPUB can be compared to each of the categories FACNPB, RESNPB, RESPUB, and ADMIN. Similarly, FACNPB can be compared with RESNPB, RESPUB, and ADMIN, and so on, for a total of ten possible comparisons. The lack of ordering, let alone a metric, among the categories of the dependent variable thus adds considerable complexity to the analysis. Rather than being able to relate each independent variable to the dependent variable by means of a single coefficient, each independent variable requires a single coefficient for each of the nonredundant comparisons.[2]

[2] With k dependent categories there are $(k^2-k)/2$ possible comparisons and thus possible coefficients for a given independent variable. However, only $k-1$ of these will contain unique information.

The multiplicative logit coefficients labeled "b*" in Table 2 indicate the factor by which the odds of one outcome versus another will change for a standard deviation change in a given variable, all other variables being held constant. Coefficients greater than one indicate an increase in the likelihood of the first outcome compared to the second outcome; coefficients less than one indicate a decrease in the relative likelihoods. For independent variables that are binary, such as having a fellowship, the unstandardized coefficient is also useful. It may be interpreted as the multiplicative change in the odds when an individual has the characteristic indicated by the dummy variable. For our purposes one of the most useful outcomes of such analyses is the ability to rank outcome categories (i.e., FACPUB, RESPUB, RESNPB, FACNPB, and ADMIN) according to how strongly each independent variable differentiates alternative outcomes. An outcome that is ranked highly will become more likely for those with high values of the independent variable. The parameters described above are shown in Table 2; the rankings and parameter magnitudes are summarized in Figure 1.

Each of the subfigures in Figure 1 can be interpreted as follows. The horizontal axis represents the magnitude of the coefficients comparing a given outcome cate-

Table 2. Multinominal Logit Analysis of Determinants of First Postfellowship Job (N = 545)

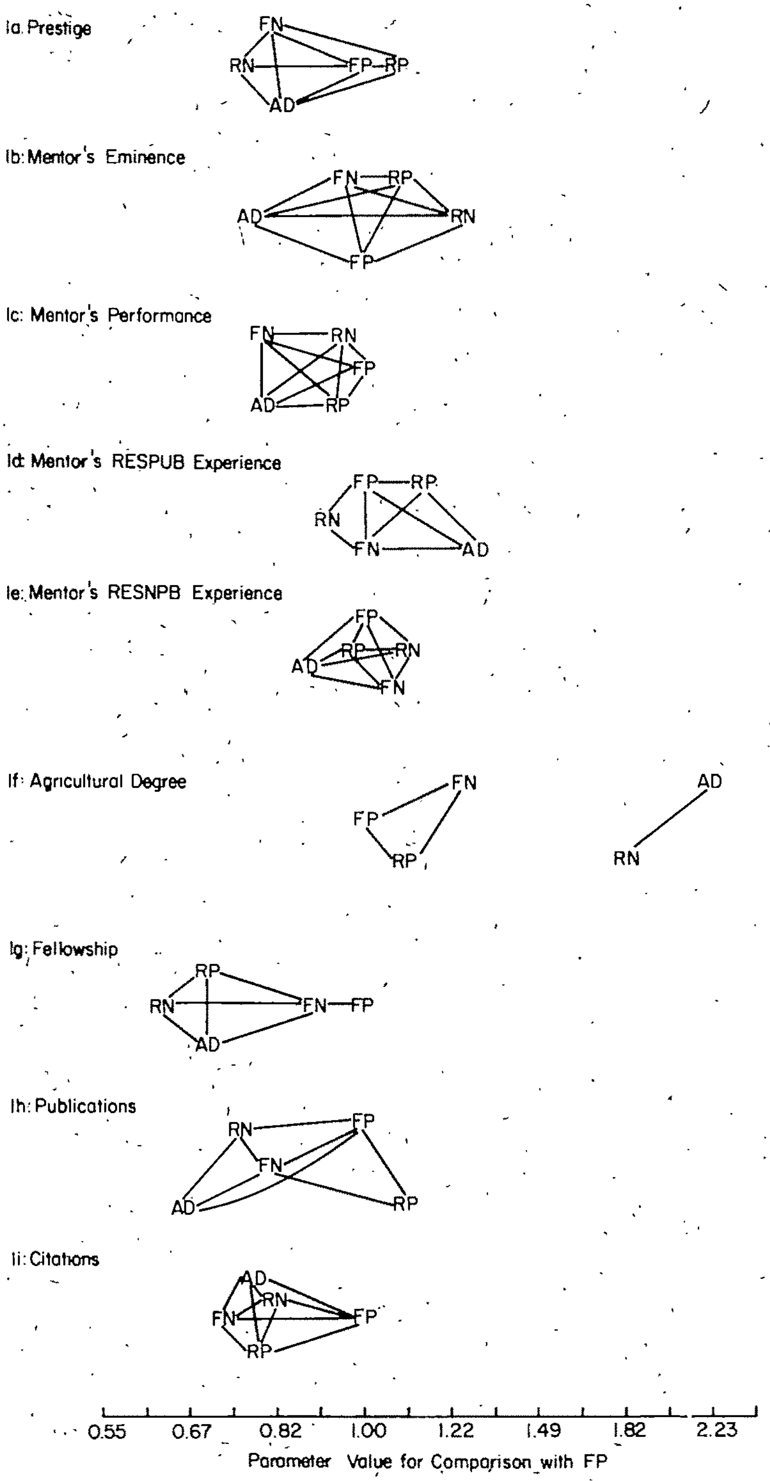
Comparison Groups		Coefficients of PRST	MPRST	MPERF	MRESPUB	MRESNPB	AGDEG	FEL	PUB	CIT
1. RESPUB/	b*	1.05	1.10	0.93	1.14	0.99	1.09	0.70**	1.11	0.76
FACPUB	b				1.31	0.99	1.20	0.50		
2. FACNPB/	b*	0.81	0.95	0.79	1.01	1.05	1.23	0.87	0.81	0.73
FACPUB	b				1.01	1.16	1.53	0.76		
3. RESNPB/	b*	0.78	1.26	0.95	0.92	1.10	1.81**	0.64**	0.75	0.80
FACPUB	b				0.84	1.29	3.48	0.40		
4. ADMIN/	b*	0.83	0.77	0.79	1.29	0.88	2.20**	0.70	0.66	0.77
FACPUB	b				1.71	0.71	5.22	0.49		
5. FACNPB/	b*	0.77	0.87	0.85	0.89	1.06	1.12	1.24	0.73	0.95
RESPUB	b				0.78	1.17	1.27	0.43		
6. RESNPB/	b*	0.74*	1.15	1.03	0.81	1.11	1.65**	0.90	0.68*	1.04
RESPUB	b				0.64	1.31	2.88	0.82		
7. ADMIN/	b*	0.79	0.71	0.85	1.14	0.89	2.02**	1.00	0.60*	1.01
RESPUB	b				1.31	0.72	4.33	1.00		
8. RESNPB/	b*	0.96	1.32	1.21	0.91	1.04	1.48*	0.73	0.93	1.09
FACNPB	b				0.83	1.12	2.27	0.53		
9. ADMIN/	b*	1.02	0.81	1.00	1.28	0.83	1.80**	0.81	0.82	1.06
FACNPB	b				1.69	0.62	3.40	0.65		
10. ADMIN/	b*	1.06	0.62	0.82	1.40	0.80	1.21	1.10	0.88	0.97
RESNPB	b				2.04	0.55	1.50	1.22		

NOTE: Item identifications are: PRST = prestige of Ph.D. department if nonfellow, prestige of fellowship department if fellow; MPRST = prestige or eminence of mentor as weighted sum of awards (see text for details); MPERF = research performance or productivity of mentor during the three-year period after the student's Ph.D., measured as square root of citations to articles in that period; MRESPUB = dummy variable with value 1 if mentor has experience in a RESPUB context; MRESNPB = dummy variable with value 1 if mentor has experience in a RESNPB context; AGDEG = 1 if Ph.D. was from department in an agricultural school, 0 if not; FEL = dummy variable with a value of 1 if postdoctoral fellowship was held, 0 if not; PUB = square root of number of publications in the three-year period ending the first year of the job; CIT = square root of the number of standardized citations to publications in the three-year period ending the first year of the job.

Row b* gives standardized logit coefficients indicating the change in the odds of the first context versus the second context for a standard deviation change in the given independent variable; row b gives unstandardized logit coefficients for dummy variables indicating the change in the odds for a unit change in the independent variable.

* Significant at .10 level, two-tailed test.

** Significant at .05 level.



KEY: FP = Faculty in a research university; FN = Faculty in a nonresearch university or college; RP = Research in a location encouraging publication; RN = Research in a location restricting publication; AD = administration. Contexts located to the right in figures are associated with higher values of the independent variable.

Figure 1. Summary of Multinomial Logit Analysis

gory to the baseline category of FACPUB, measured on a logarithmic scale.[3] Thus, the further two categories are apart on the horizontal axis, the more strongly a standard deviation change in the independent variable differentiates them. If one category is to the right of a second category in the diagram, the odds of the first category are increased relative to the second category by an increase in the magnitude of the independent variable. All subfigures are drawn in the same scale and accordingly the relative distances between categories from one subfigure can be compared to those in any other subfigure. The distances on the vertical axis have no substantive interpretation, being added solely to facilitate the drawing of the figures. Finally, the lines between categories indicate the lack of a statistically significant difference (in terms of the given independent variable) between two categories. Thus, categories that are connected by a line are linked in the sense that the given independent variable does not significantly differentiate the likelihood of one outcome occurring compared to another. While general orderings were expected, these analyses are exploratory given the lack of prior research in this area. Accordingly, the test of significance represented by the lines are two-tailed tests at the .10 level of significance. Had one-tailed tests based on the suspected orderings been employed, somewhat stronger differentiation among categories would have been obtained. In evaluating the results presented in Figure 1, it is important to consider not only specific instances of statistical significance, but also the overall patterns of effects of a given variable.

Having a prestigious origin is associated with accepting a job in a location which encourages publication. Such locations, which are generally considered to be the most prestigious, include faculty positions in research universities and research positions in Federally Funded Research and Development Laboratories, private nonprofit laboratories, and academic research centers (FACPUB and RESPUB). An increase of one standard deviation in prestige, roughly the difference between a distinguished and a strong location on the Cartter prestige scale, increases the odds of such positions over the contexts of administration, college teaching, and industrial research by an average factor of 1.27 (the inverse of the average of the effects in Equations 2 through 7 in Table 2). Thus, those from distinguished locations are 2.05 (= 1.27^3) times as likely to accept a position which encourages publication over any other type of job than their counterparts who obtained degrees in adequate-plus departments, all other variables being held constant. As shown graphically in Figure 1a, the prestige of the doctoral or fellowship department tends to differentiate organizational contexts into two groups. The first consists of faculty positions in research universities or research positions in locations which encourage publication; the second consists of contexts where some sort of structural constraint on publishing is often found.

Overall the effects of characteristics of the mentor on the organizational context of the student's first job are quite small. There are, however, several interesting configurations of effects. While prestige of the origin department and our measure of the mentor's eminence are significantly correlated (r = .28), the more eminent the mentor, the more likely a position where research is the primary role, regardless of encouragement or opportunity for publication. Thus, a prestigious departmental origin will decrease the chances of an industrial research position, while having an eminent mentor will increase the chances. Effects of the mentor's current research performance are not statistically significant, but nonetheless demonstrate a consistent pattern of differentiating research positions from nonresearch positions. Faculty positions in research universities and research positions in any sector are

[3] The following relationship holds among additive parameters in a multinomial logit analysis, where d_{ij} is the additive logit coefficient comparing dependent categories i and j: $d_{ij} = d_{ik} - d_{jk}$. Hence, the relative distances among categories can be plotted in a one-dimensional space with no loss of information. Since we are using multiplicative parameters, which are simply the exponential transformations of the additive parameters, the coefficients must be plotted on a logarithmic scale.

positively associated with the mentor's research productivity.

The types of work experiences had by the mentor may affect the type of job obtained by the student. Accordingly, two dummy variables were coded indicating whether the mentor had work experience in either a RESPUB or RESNPB location. While the effects of these variables are not statistically significant, the most likely work activity of a student having a mentor with work experience in an industrial research position (RESNPB) is an industrial research position. The most common organizational context of those having mentors with nonfaculty, academic research experience (RESPUB) is either administration (which could be in a lab encouraging publication) or research in a location encouraging publication.

Having an agricultural degree is the strongest factor affecting choice of context, even after controlling for its effect on obtaining a fellowship. Those with degrees from agricultural schools are on average nearly five times as likely to accept administrative positions and over three times as likely to accept industrial research positions than positions encouraging publication (Equations 3, 4, 6, and 7). Those from agricultural schools are also significantly less likely to accept teaching positions in nonresearch universities and colleges (Equations 8 and 9). As shown in Figure 1f, whether or not a degree is obtained in an agricultural school divides organizational contexts into two distinct groups. The first group, which is associated with an agricultural degree, consists of primarily industrial positions. The second group consists of nonindustrial contexts, with teaching in a research university or doing research in a location encouraging publication being least associated with an agricultural degree. While it is impossible to determine the exact process leading to these outcomes, they suggest both a greater growth in nonagricultural biochemistry in the universities, resulting in more teaching positions for nonagricultural biochemists than agricultural biochemists, and a greater demand for agricultural biochemists in the food insutry.

Postdoctoral fellows are nearly twice as likely to obtain faculty positions in research universities (FACPUB) as jobs in any other context (Equations 1 to 4), all other factors being held constant. The fellowship does not significantly differentiate among the remaining contexts, although college teaching is more likely than any of the nonfaculty contexts that restrict publication. Overall, postdoctoral fellowships lead to academic positions, primarily in research universities.

Citations at the time the first job is obtained have no effects on the likelihood of various job alternatives (see Figure 1i). The number of citations to a student's pre-employment publications increases the odds of a faculty position in a research university by a factor of 1.3, although this tendency is not statistically significant.

Number of publications has a stronger effect, differentiating organizational contexts which encourage or demand publication from those which explicity or implicitly restrict publication. Figure 1h shows that industrial research jobs and administrative positions are associated with few publications, jobs in research universities and locations encouraging publication (RESPUB and FACPUB) are associated with more publications, and college teaching falls in between the two other groups.

The fact that citations have no effects at all in the allocation of positions, while number of publications does differentiate those positions which strongly encourage publication from those which discourage publication, suggests that scientists who have published and want to continue publishing seek out and obtain positions where they will be rewarded for an activity at which they have already been successful, as opposed to positions where publication is simply tolerated. Past studies which have reported that the "best" students go to nonindustrial locations may have equated "best" with "most publications," regardless of the quality of those publications.

Overall our findings suggest several things. Productivity is important only in terms of quantity of publications. Those who publish before obtaining a permanent job have a tendency to obtain jobs which encourage publication. The nature of the

graduate education and the postdoctoral fellowship have a variety of effects. Prestige is associated with accepting a faculty position in a research university and with not obtaining an industrial research position. Having an eminent mentor, on the other hand, is associated with accepting a research position as opposed to a faculty position. Beginning the career with a postdoctoral fellowship increases the chances of obtaining a faculty position in a research university.

In assessing these findings it is important to realize that these are partial effects which remain after statistically controlling for all other variables in the logit analysis. Thus, since fellows tend to increase their origin prestige by replacing the prestige of the doctoral department by that of the postdoctoral location and to increase their level of productivity during the period of the fellowship, and since both of these changes will increase the odds of a faculty position in a research university over an industrial research position, the total effect of the fellowship is greater than the direct effect of the dummy fellowship variable. Further, while a nonagricultural degree directly increases the likelihood of a faculty position in a research university, it also increases the likelihood through its effect on the acceptance of a postdoctoral fellowship.

THE EFFECTS OF ORGANIZATIONAL CONTEXT ON PRODUCTIVITY

To control for the effects of changes in organizational context on productivity, our analysis is initially restricted to those scientists who remain in the same job for at least nine years. This period of time should be long enough to allow the effects of organizational context to emerge. To assess the effects of changes in organizational contexts and to determine how responsive a scientist's productivity is to a change in his work environment, those scientists who change contexts are then examined.

The effects of organizational context are introduced into regressions on productivity by means of a set of four dummy variables, one coefficient for each of the categories of administration (ADMIN), research in a location encouraging publication (RESPUB), research in a location limiting publication (RESNPB), and teaching in a college (FACNPB). Having a faculty position in a research university (FACPUB) is the fifth, excluded category that serves as a standard of comparision for the other contexts.

After three years in the job, significant effects of context on level of publication begin to appear (Equation 1, Table 3). Having a teaching position in a college (FACNPB) decreases the level of publication relative to faculty in research universities by an average of .56 on the transformed scale of publications, after controlling for characteristics of the doctoral and/or postdoctoral experience.[4] The effect of an industrial research position (RESNPB) is a relative decrease of .74 publications. There are no significant differences in publication levels between faculty in research universities and the two remaining categories.

To take into account the initial differences in productivity among scientists entering different organizational contexts, Equation 2 controls for the productivity level of a scientist at the time he begins his job. In such an equation, the effect of the contextual variables can be thought of as the change in productivity that has occurred after a scientist has entered a particular organizational context. In short, the effects of differing levels of productivity upon entering the job have been removed. The strongest effect is now stability. A scientist who publishes before entering a particular context tends to continue publishing, independently of context. Still, organizational context remains a significant factor, albeit with attenuated effects. Working in industrial research (RESNPB) has the strongest depressing contextual effect with an unstandardized coefficient of $-.49$; the second largest effect is for college teaching with a coeffi-

[4] A square root transformation has been made for both publications and citations; thus the differences between adjacent productivity counts (e.g., five and six publications; nine and ten citations) decrease as the level of productivity increases. In this particular case the use of the transformation implies that contextual effects are stronger in terms of raw publications among the more productive scientists in a location.

Table 3. Determinants of Publications for Those Staying at the Same Job for at Least Nine Years (N = 308)

Publication Year		Coefficients of PRST	MPRST	MPERF	AGDEG	FEL	RESPUB	FACNPB	RESNPB	ADMIN	(From Year One) PUB	(From Year One) CIT	R^2
1. Year 3	b	.0343	−.00488	.0390*	−.136	.273**	.0620	−.559**	−.741**	−.141			.143
	b^s	.035	−.021	.101	−.066	.136	.024	−.159	−.243	−.041			
2. Year 3	b	.0407	.00187	−.0113	−.0272	−.00328	.132	−.346**	−.493**	−.0157	.615**	.0364	.460
	b^s	.041	.008	−.029	−.013	−.002	.051	−.098	−.162	−.004	.560	.074	
	r	.116	.049	.155	−.146	.193	.063	−.149	−.253	−.045	.646	.498	
3. Year 6	b	.104*	.000250	.0469**	−.0498	.283**	−.0233	−.961**	−.936**	−.473**			.193
	b^s	.090	.001	.104	−.021	.121	−.008	−.234	−.263	−.117			
4. Year 6	b	.109**	.00558	.00284	.0533	.600	.0419**	−.802**	−.755**	−.408**	.391**	.0714**	.325
	b^s	.094	.021	.006	.022	.026	.014	−.195	−.212	−.101	.305	.125	
	r	.199	.093	.210	−.108	.200	.053	−.220	−.239	−.104	.472	.410	
5. Year 9	b	.165**	−.00782	.0525**	−.103	.283**	.0508	−.885**	−.961**	−.391**			.206
	b^s	.139	−.208	.114	−.042	.118	.016	−.210	−.264	−.095			
6. Year 9	b	.172**	−.00199	−.000956	.0310	.0319	.134	−.722**	−.778**	−.360**	.316**	.129**	.350
	b^s	.145	−.007	−.002	.013	.013	.043	−.172	−.214	−.087	.241	.221	
	r	.236	.077	.225	−.118	.201	.071	−.209	−.253	−.093	.473	.455	
											(From Year Six) PUB	(From Year Six) CIT	
7. Year 9	b	.0779*	−.00834	.0182	−.0341	.0276	.0792	−.240*	−.367**	−.0671	.582**	.0860**	.592
	b^s	.066	−.030	.040	−.014	.012	.026	−.056	−.101	−.016	.569	.164	
	r	.236	.077	.225	−.118	.201	.071	−.209	−.253	−.093	.747	.612	

NOTE: Item identifications are: RESPUB = 1 if organizational context is RESPUB, 0 if not; FACNPB = 1 if context is FACNPB, 0 if not; RESNPB = 1 if context is RESNPB, 0 if not; ADMIN = 1 if context is ADMIN, 0 if not; FACPUB is the excluded context; see Table 2 for other item identifications.

Row b gives unstandardized regression coefficients; row b^s gives standardized regression coefficients; row r gives zero order correlations with dependent variable.

* Significant at .10 level, one-tailed test.

** Significant at .05 level.

cient of −.35. The effect of administration remains small and insignificant; the effect of research in a location that encourages publication is positive (although not significant at the .10 level) and of increased magnitude, suggesting a slight increase in the productivity levels of these scientists.

The effects of characteristics of the doctoral and postdoctoral experience change significantly when change in the level of publication is examined (Equation 2), rather than the level of publication (Equation 1). While having a fellowship significantly increases the *level* of publication, adding .27 transformed publications, it has no effect on *changes* in the level of publication which occur after the job is obtained. Thus, the experience of the fellowship does not directly affect subsequent levels of productivity. Rather, its effect is indirect, operating via its positive effect on productivity during the period of the fellowship and its effect on the allocation of jobs. Other variables related to the doctorate and fellowship have no significant effects on either level of or change in publication, with the exception of the performance of the mentor, which has a barely significant positive effect. In regressions on level of publication excluding contextual variables (regressions not shown), the effects of all of the independent variables were stronger with increased levels of statistical significance. It appears that most of the effects of these characteristics occur in the form of cumulative advantage. They affect early productivity and initial job placement, which in turn affect later productivity, independently of their own determinants.

The effects of educational characteristics remain basically unchanged at the sixth year (Equation 3), with an increase in the magnitude and significance of the effect of origin prestige on both level and change in publications, and an increased effect of the mentor's performance on level of publication. The effects of context on level of publication in the three-year period ending in year six of the job are similar to those found in year three, with the exception of the expected emergence of the negative effect of administration on the level of publication.

The prestige of the origin has a positive effect on the change in publication level over the first six years of the job. Simply having a fellowship, attending an agricultural school, or having an eminent or productive mentor do not. The strength of the contextual effects has increased. Compared to those with faculty positions in research universities, those in administration have significantly decreased their rate of publication by .51 transformed articles over a three-year period. Such a decrease is even stronger for those in colleges and industrial research, with decreases of nearly one article. There are no significant differences between faculty in research universities and researchers in locations encouraging publication. Finally, as before, there is a strong effect of the past level of publication on the present level of publication, although the magnitude has decreased slightly, as would be expected due to the greater lag between productivity measures. Past citations are beginning to have a positive effect on changes in productivity, suggesting the emergence of a reinforcing effect of recognition on publication.

By the ninth year the effects on level of publication of having a prestigious origin or a productive mentor have increased, while the effect of having a fellowship is unchanged (Equation 5). The effects of context on the level of publication are essentially unchanged from year six. Equation 6 continues the trend of effects on changes in level of productivity that appeared in Equation 4.

To assess the relative strengths of the effects of earlier productivity and organizational context, a sheaf coefficient (Heise, 1972) for the contextual variables can be computed. Its value for Equation 6 is 0.28, indicating a stronger overall effect of context than either number of publications or number of citations, and a slightly weaker effect than the combined effects of early citations and publications. Productivity at the start of the job remains a strong factor predicting future productivity, although its effect decreases with time, while the effect of context increases.

The nature of the effect of organizational context on productivity is further illustrated in Equation 7 of Table 3 in which level of publication in year 9 is re-

gressed on the same variables as in Equation 6, with the substitution of levels of productivity from year six for productivity in year one. In this equation the effects of the contextual variables on changes in level of publication are of only slightly decreased magnitudes. Thus, organizational context has a continuing effect on level of publication lasting at least nine years into the job.

Similar analyses are possible with the number of citations received by scientists in various contexts, where citations can be thought of as an indicator of the recognition that published work has received (see Table 4). Results are similar to those obtained for publications.

As with publications, the productivity measures in Table 4 are for the three-year period ending in a given year of the job. How many years this is from the year of the doctorate depends on whether or not a postdoctoral fellowship was held before the first job. Thus, the citation measure for year six is computed from citations to articles in the fourth through sixth years of the job. Citations were coded from a volume of *Science Citation Index* published one or two years after the sixth year. Complete details on the construction and standardization of these measures are presented in Long (1978).

While the contextual effects are similar, although taking somewhat longer to emerge, the effects of background characteristics differ. For all years examined (see Equations 1, 3, and 5), an agricultural degree significantly decreases the number of citations received, a prestigious origin increasingly increases citations, and a productive mentor increases citations. While having a fellowship has a strong positive effect on citations in the third year, the effect diminishes by the sixth year and turns negative by year nine. Thus, even though the prestige of the fellowship along with the prestige of the doctorate continue to positively affect the level of productivity, simply having a fellowship does not maintain its positive effect after controlling for other factors.

By the third year in a given context, there is one significant contextual effect on the level of citations—the negative effect of more than one transformed citation for those who hold industrial research positions compared to faculty in research universities. By year nine there are also significant negative effects for faculty in colleges and administrators. Finally, those with research positions in nonindustrial locations (RESPUB) are not significantly differentiated from faculty in research universities.

Equations 2, 4, and 6 examine the effects of educational characteristics and contextual effects on changes in levels of citation. In each equation productivity at the first year of the job is controlled. By year six negative and statistically significant effects of administration, college teaching, and industrial research are found; by year nine these effects have increased, with a sheaf coefficient of 0.20. Hence, not only are levels of citation negatively affected by context, but so are the changes in citations received. As with publications, it takes time before the contextual effects emerge. Equation 7 shows that the negative effects continue, affecting the change in citations from the sixth to the ninth year in the job.

When changes in level of citation are examined, effects of educational variables change. Having an agricultural degree no longer has a significant effect, nor does the eminence of the mentor. Having a prestigious fellowship or doctoral department, however, strongly increases the change in level of productivity. Simply having a fellowship has an initial positive effect which becomes negative by year nine.

Organizational context emerges as a strong factor determining not only levels of productivity, but also, and more importantly, changes in rates of productivity occurring after a position is obtained. Administrators, faculty in colleges, and researchers in locations limiting publication decrease their productivity relative to faculty and researchers in research universities. These contextual effects go beyond initial differences in doctoral education and pre-employment productivity, which were shown to affect the allocation of positions. For after controlling for these variables, contextual effects remain strong and statistically significant. Accordingly, it is not that those in research

Table 4. Determinants of Citations for Those Staying at the Same Job for at Least Nine Years (N = 308)

Citation Year		Coefficients of: PRST	MPRST	MPERF	AGDEG	FEL	RESPUB	FACNPB	RESNPB	ADMIN	(From Year One) PUB	(From Year One) CIT	R^2
1. Year 3	b	.191*	−.0124	.129**	−.666**	1.33**	.266	−.307	−1.05**	.197			.245
	b^s	.090	−.025	.156	−.151	.311	.048	−.041	−.161	.027			
2. Year 3	b	.210**	.00166	−.0331	−.208	.687**	.546**	.014	−.713**	.00933	.0387	.642**	.541
	b^s	.099	.003	−.040	−.047	.160	.098	.002	−.109	.001	.016	.613	
	r	.197	.103	.232	−.240	.380	.015	−.066	−.229	−.049	.522	.690	
3. Year 6	b	.309**	.00437	.0811*	−.463**	1.05**	−.172	−.988**	−.566*	−.564*			.177
	b^s	.137	.008	.092	−.099	.231	−.029	−.123	−.082	−.072			
4. Year 6	b	.323**	.0158	−.0292	−.179	.554**	.00432	−.679**	−.223	−.545*	.508**	.306**	.330
	b^s	.143	.030	−.033	−.038	.121	.001	−.085	−.032	−.069	.203	.275	
	r	.238	.132	.220	−.163	.312	−.037	−.138	−.096	−.116	.470	.496	
5. Year 9	b	.408**	.00932	.115**	−.695**	−.814**	−.132	−1.28**	−1.21**	−.761**			.159
	b^s	.176	.017	.127	−.144	−.174	−.022	−.155	−.170	−.094			
6. Year 9	b	.420**	.0185	.0184	−.434*	−1.22**	.0282	−1.05**	−.966**	−.810**	.228	.326**	.255
	b^s	.181	.034	.020	−.090	−.261	.005	−.127	−.136	−.100	.088	.285	
	r	.220	.105	.239	−.142	−.059	.051	−.167	−.144	−.088	.306	.356	
											(From Year Six) PUB	(From Year Six) CIT	
7. Year 9	b	.241**	.00739	.0638*	−.482**	−1.36**	−.0514	−.514*	−.637**	−.352	.363**	.419**	.405
	b^s	.104	.014	.071	−.100	−.290	−.008	−.062	−.089	−.043	.181	.408	
	r	.220	.105	.239	−.142	−.059	.051	−.167	−.144	−.088	.500	.527	

NOTE: See Tables 2 and 3 for item identifications.

Row b gives unstandardized regression coefficients; row b^s gives standardized regression coefficients; row r gives zero order correlations with dependent variable.

* Significant at .10 level, one-tailed test.

** Significant at .05 level.

universities are more capable, as indicated by previous levels of publication and citation, or better trained, as indicated by characteristics of their graduate education, than their counterparts in other contexts. The fact that it takes time for the contextual effects to emerge suggests that these effects are not simply the result of immediate changes in scientists' goals or categorical barriers to publication in some contexts. Scientists in administration, industry, and colleges begin their employment in those locations with continued levels of productivity, but in time the effects of earlier productivity decrease while the effects of context increase.

PRODUCTIVITY, ORGANIZATIONAL CONTEXT, AND MOBILITY

Changes in organizational context are relatively infrequent. For the 557 scientists in this study, only 275 changes in context (as defined here) involving 208 individuals occurred. Thus, for the 6915 career-years for which data are available, changes occur on an average of once every 25 years, affecting 37% of the sample. A majority of these changes are in roles rather than sectors, such as moving into academic administration from a research or faculty position in academia, or moving from research to administration in industry.

Of the 275 moves, 69 were selected for analysis. To ensure independence of observations in the regression analyses, multiple moves by the same individual were excluded, reducing the number of possible moves to 208. Of the remaining moves, 44% were excluded because of censoring on the second job before it lasted six years; in such cases no information was available regarding the scientists' eventual duration in the second job. To allow for compatability with the analyses of the stayers, 47 moves were excluded that were not censored, but which had a first job lasting less than three years or a second job lasting less than six years. Of these, 47% involved jobs lasting one year. The remaining sample included 69 moves. Analyses were also run with the 94 job pairs that had at least two years of data for each job. Results were similar, although as would be expected by the limited amount of time for contextual effects to operate, the magnitudes of the effects were attenuated. Given the small size of the sample, the results must be viewed with consequent caution.

Equation 1 of Table 5 regresses publications at the time of the job change on educational characteristics, the *first* organizational context, and measures of publications and citations three years prior to the move.[5] Except for the barely significant, negative effect of the mentor's eminence, the effects of educational characteristics are not significant. The effects of contextual locations are in the expected directions and are significant at the .10 or higher level. As for the stayers, the effect of research in an industrial location is large, as is the effect for teaching in a college. Unlike the case for stayers, there are significant differences between faculty at research universities and researchers in such locations. Finally, the effect of past levels of productivity on current levels of publication are positive. The effect of publications is strongest and significant at the .01 level; the effect of citations is significant and approximately two-thirds the size of the publication effect.

Equation 2 presents comparable results for the sixth year in the second organizational context. The independent variables now include indicators of the *second* organizational context and levels of productivity in the *three-year period before* the change of contexts. The similarities between Equation 2 of Table 5 and Equation 4 of Table 3 for stayers are important, as is the one major difference. Educational characteristics do not have significant effects (which in part reflects the small sample size). Organizational contexts have effects in the same directions and of roughly similar magnitudes. Interestingly, the effects of past productivity on current productivity are greatly at-

[5] Three-year publication counts for the dependent variable are for the period ending the first year in the second job. This lag of one year is introduced to allow for the time between the submission of manuscripts and their publication. Because these scientists had their first job for an average of 5.1 years, the results would be expected to fall between those presented in Equations 2 and 4 of Table 3 and 4.

Table 5. Determinants of Publications for Movers Staying in the Second Job at Least Six Years (N = 69)

Publication Year		Coefficients of											
								(Context of First Job)			(Three-Year Period Ending Two Years Before Move)		
		PRST	MPRST	MPERF	AGDEG	FEL	RESPUB	FACNPB	RESNPB	ADMIN	PUB	CIT	R^2
1. Year 1	b	.015	−.0712*	.0108	−.226	−.321	−.560**	−1.32**	−1.14**	−.664*	.406**	.114**	.507
in 2nd	b*	.014	−.204	.023	−.110	−.137	−.266	−.299	−.512	−.207	.384	.265	
Context	r	.070	−.150	.150	−.226	.077	.216	−.007	−.431	.050	.582	.487	
								(Context of Second Job)			(Three-Year Period Ending One Year After Move)		
2. Year 6	b	.120	.0291	−.0278	−.0943	.244	.180	−1.33**	−.567*	−.491**	.158	.00925	.264
in 2nd	b*	.118	.087	−.065	−.048	.109	.059	−.278	−.185	−.244	.165	.020	
Context	r	.213	.128	.127	−.170	.256	.207	−.215	−.140	−.241	.223	.295	

NOTE: See Tables 2 and 3 for item identifications.

Row b gives unstandardized regression coefficients; row b* gives standardized regression coefficients; row r gives zero order correlations with dependent variable.

* Significant at .10 level, one-tailed test.

** Significant at .05 level.

tenuated. That is, the productivity that resulted from work in the first organizational context has a much smaller effect on level of publication in the current context. This differs both from the results in Equation 1 of Table 5 and Equation 4 of Table 3 for the stayers. Within the same context, past productivity influences current productivity. Once mobility occurs across sectors, only organizational context significantly affects publication levels. Past organizational context of employment has no further effect on productivity in the new context (regression not shown). This lack of effect further underscores the potency of current organizational context as a determinant of productivity.

Equation 1 of Table 6 shows that educational characteristics have no significant effects on changes in the citation level within the first context. Moderate effects of organizational context are found in the expected directions, but are generally significant at only about the .20 level. Industrial research is the one exception, being significant at the .05 level. Finally, past levels of citation influence current levels, with the standardized coefficient being large (b^s=.56) and significant at the .01 level.

In Equation 2 the results are presented for the sixth year in the *second* context. As with publications, productivity during the last three years in the *first* context has a smaller effect than for stayers. Effects of contextual location are in the expected direction. As with publications, the effects of the first context are very small and not nearly significant (regression not shown). One interesting difference is that educational characteristics emerge as significant factors affecting changes in citation level. Having an eminent mentor positively affects changes in citation level at the new job. Obtaining a degree from an agricultural school has a strong negative effect.

The major finding of this section is that when changes of organizational context occur, changes in productivity follow. These changes are not strongly affected by past productivity or past organizational context. Rather, the organizational context of the new position quickly emerges as the significant factor influencing productivity. Since contextual effects on levels of productivity in the movers' first context conform to the results obtained for the stayers, it is unlikely that the movers represent those who were unsuccessful in meeting the expectations of their first job context. The argument that the most productive scientists end up in the most prestigious locations is not supported by our data, since, when mobility occurs, the once-productive scientists suffer declines in their productivity—declines that are determined by the new organizational context. The argument would hold only if one were willing to assume that a scientist's ability to make contributions to the body of scientific knowledge changes along with a change of organizational context. This, however, seems unlikely.

SUMMARY AND CONCLUSIONS

In an earlier study of the academic career, Long (1978) examined the relationship between a scientist's productivity and the prestige of his academic position. He found that while a scientist's productivity does not affect the prestige of the academic position obtained, the prestige of the position does affect the scientist's later productivity. Those in prestigious locations increased their current productivity independently of earlier productivity. Our current study shows that scientific productivity and organizational context are similarly related. The chances of obtaining employment in a particular organizational context are not strongly affected by productivity as indicated by either published papers or the scientific community's utilization of those papers. While there is some tendency for those who have published to obtain positions in contexts which encourage publication, no similar effect is found when citations are considered. Within three to six years of obtaining a position in a specific context, a scientist's level of productivity conforms with the characteristics of that context, independently of previous productivity. Compared to those in research universities, those doing industrial research, teaching in four-year colleges, or administering a laboratory decline in the

Table 6. Determinants of Citations for Movers Staying in the Second Job at Least Six Years (N = 69)

Citation Year		Coefficients of											
								(Context of First Job)			(Three-Year Period Ending Two Years Before Move)		
		PRST	MPRST	MPERF	AGDEG	FEL	RESPUB	FACNPB	RESNPB	ADMIN	PUB	CIT	R^2
1. Year 1	b	.181	−.0501	−.0167	.183	.259	−.498	−1.33	−1.20**	−.709	.192	.488**	.507
in 2nd	b^s	.084	−.071	−.018	.044	.055	−.117	−.149	−.265	−.109	.090	.563	
Context	r	.212	.058	.228	−.305	.327	.263	−.074	−.344	.034	.524	.672	
								(Context of Second Job)			(Three-Year Period Ending One Year After Move)		
2. Year 6	b	.140	.215*	.0391	−1.55**	.509	−.626	−2.63**	−1.02	−1.20**	−.0534	.265	.420
in 2nd	b^s	.051	.237	.033	−.292	.084	−.075	−.201	−.112	−.220	−.021	.207	
Context	r	.272	.343	.276	−.397	.300	.070	−.161	−.151	−.228	.210	.403	

NOTE: See Tables 2 and 3 for item identifications.

Row b gives unstandardized regression coefficients; row b^s gives standardized regression coefficients; row r gives zero order correlations with dependent variable.

* Significant at .10 level, one-tailed test.

** Significant at .05 level.

number of and citations to their publications. The importance of context is further demonstrated by the flexibility in productivity patterns which is observed among those who move from one context of work to another. Within six years the levels of productivity of the movers have changed to the extent that past productivity no longer predicts future productivity. New levels of productivity are determined largely by the context of work.

These results do not support the idea that scientists are allocated to various contexts on the basis of contributions to the body of scientific knowledge. Past research indicating that the most productive scientists are recruited to academic locations may have confused the *cause* of a scientist working in a given organizational context with the *effect* of working in that context.

Our findings provide additional evidence for the operation of cumulative advantage in the scientific career (cf. Allison and Stewart, 1974; Cole and Cole, 1973): initial advantages are soon transformed into additional advantages. For example, having a prestigious doctoral department increases the chances of the graduate obtaining a postdoctoral fellowship or a position in a research university. The position in a research university results in increased productivity; the fellowship directly increases publications over the fellowship period and indirectly increases productivity by increasing the chances of obtaining a faculty position in a research university, thus providing a secondary effect of doctoral prestige. The fact that the major determinants of productivity a decade into the career are not direct effects of education and training, but largely the effects of past productivity and organizational context of employment, points to the extreme importance of the recruitment process in explaining inequalities among scientists in productivity. That recruitment operates independently of productivity suggests that, to the extent that publications and citations reflect a scientist's contribution to scientific knowledge, the stratification system in science is far from universalistic. Rather, initial advantages are accumulated, resulting in the well-known inequality among scientists in levels of productivity.

REFERENCES

Allison, Paul D.
1976 "Processes of stratification in science." Ph.D. Dissertation. University of Wisconsin, Madison.

Allison, Paul D. and John A. Stewart
1974 "Productivity differences among scientists." American Sociological Review 39:596–606.

American Chemical Society
Directory of Graduate Research. Washington, D.C.: American Chemical Society.

Bernal, J. D.
[1954] 1971 Science in History. Volume 3: The Natural Sciences in Our Time. Boston: M.I.T. Press.

Caplow, Theodore and Reece McGee
1958 The Academic Marketplace. Garden City, New York: Doubleday.

Cartter, Allan M.
1966 An Assessment of Quality in Graduate Education. Washington: American Council on Education.

Cattel Press
American Men and Women of Science. New York: R. R. Bowker.

Chemical Abstracts Service
Chemical Abstracts. Columbus, Ohio: American Chemical Society.

Cole, Jonathan R. and Steven Cole
1973 Social Stratification in Science. Chicago: University of Chicago Press.

Cotgrove, Stephen and Steven Box
1970 Science, Industry and Society. London: Allen and Unwin.

Curtis, Richard B.
1969 The Invisible University. Washington: National Academy of Sciences–National Research Council.

Domencich, T. and D. McFadden
1975 Urban Travel Demand. Amsterdam: North-Holland.

Gaston, Jerry
1978 The Reward System in British and American Science. New York: Wiley.

Glaser, Barney G.
1964 Organizational Scientists. New York: Bobbs-Merrill.

Goodman, Leo A.
1971 "The analysis of multidimensional contingency tables." Technometrics 13:33–61.
1972 "A modified multiple regression approach to the analysis of dichotomous variables." American Sociological Review 37:28–46.

Hagstrom, Warren O.
1965 The Scientific Community. New York: Basic Books.
1971 "Inputs, outputs, and the prestige of university science departments." Sociology of Education 44:375–97.

Hargens, Lowell L. and Warren O. Hagstrom
1981 "Consensus and status attainment patterns in scientific disciplines." Working Paper. Indiana University.

Harmon, Lindsey R.
1965 Profiles of Ph.D.s in the Sciences. Wash-

ington: National Academy of Sciences–National Research Council.
1968 Careers of Ph.D.s: Academic versus Nonacademic. Washington: National Academy of Sciences–National Research Council.

Heise, David
1972 "Employing nominal variables, induced variables, and block variables in path analyses." Sociological Methods and Research 1:147–73.

Institute for Scientific Information
Science Citation Index. Philadelphia: Institute for Scientific Information.

Klaw, Spencer
1968 The New Brahmins. New York: Morrow.

Kornhauser, William
1962 Scientists in Industry. Berkeley, Ca.: University of California Press.

Krohn, R. G.
1971 The Social Shaping of Science. Westport, Conn.: Greenwood Publishing.

Long, J. Scott
1978 "Productivity and academic position in the scientific career." American Sociological Review 43:889–908.

Long, J. Scott, Paul D. Allison, and Robert McGinnis
1979 "Entrance into the academic career." American Sociological Review 44:816–30.

Long, J. Scott and Robert McGinnis
1981 "The effects of the mentor on the academic career." Working Paper. Washington State University.

Long, J. Scott, Robert McGinnis, and Paul D. Allison
1980 "The problem of junior-authored papers in constructing citation counts." Social Studies of Science 10:127–43.

Marcson, Simon
1960 The Scientists in American Industry. New York: Harper.
1972 "Research settings." Pp. 161–91 in S. Z. Nagi and R. G. Corwin (eds.), The Social Contexts of Research. London: Wiley.

McGinnis, Robert, J. Scott Long, and Paul D. Allison
1981 "Postdoctoral training in bioscience: allocation and outcomes." Social Forces, forthcoming.

Merton, Robert K.
1973 The Sociology of Science. Chicago: University of Chicago Press.

National Academy of Sciences
1974 Postdoctoral Training in the Biomedical Sciences. Washington, D.C.: National Academy of Sciences.

Nerlove, Marc and S. J. Press
1973 "Univariate and multivariate log-linear ad logistic models." RAND Report No. R-1306-EDA/NIH.

Pelz, Donald C. and Frank M. Andrews
1966 Scientists in Organizations. New York: Wiley.

Reskin, Barbara F.
1976 "Sex and status attainment in science." American Sociological Review 41:597–612.
1977 "Scientific productivity and the reward structure of science." American Sociological Review 42:491–504.
1979 "Academic sponsorship and scientists' careers." Sociology of Education 52:129–46.

Roose, Kenneth D. and Charles J. Andersen
1970 A Rating of Graduate Programs. Washington: American Council on Education.

Sklair, L.
1972 "The political sociology of science." Pp. 43–59 in P. Halmos (ed.), The Sociology of Science. The Sociological Review Monograph 18 (September).

Storer, Norman
1966 The Social System of Science. New York: Holt.

Zuckerman, Harriet
1970 "Stratification in American science." Pp. 235–57 in E. O. Laumann (ed.), Social Stratification. Indianapolis: Bobbs-Merrill.

Zuckerman, Harriet and Robert K. Merton
1971 "Patterns of evaluation in science: institutionalization, structure and functions of the referee system." Minerva 9:66–100.
1972 "Age, aging and age structure in science." Pp. 292–356 in M. W. Riley et al. (eds.), Aging and Society: Vol. III. New York: Russell Sage Foundation.

TRENDS IN THE RELATIONSHIP BETWEEN SEX AND PSYCHOLOGICAL DISTRESS: 1957–1976*

RONALD C. KESSLER
JAMES A. MCRAE, JR.
University of Michigan

In recent years a great deal of interest has developed in the "sex role hypothesis" as a way of understanding the high rates of psychological distress among women in our society. Although this work has been almost entirely cross-sectional, the underlying hypothesis predicts that the relationship between sex and distress should decline as sex roles become more comparable. Basing our analysis on three national surveys and two community surveys spanning the years 1957 to 1976—a period of rapid changes in the roles of women—we document a reduction in the relationship between sex and one indicator of distress, a screening scale of psychophysiological symptoms. Specification analyses show that the increased labor-force participation of women has been responsible for part of this trend. However, there appear to be no relationships between the decline in psychophysiological symptoms and changes in educational attainment, rates of marriage, marital dissolution, or childbearing.

Over the past two decades substantial changes in female roles have taken place. Since 1957, when the baby boom peaked, the number of women having children has steadily decreased, as has the size of the typical family (National Center for Health Statistics, 1951, 1969, 1979). Women who are having children report at increasing rates that the births were planned (David and Baldwin, 1979). This trend reflects the increased use of birth control devices, changing norms about ideal family size (U.S. Bureau of the Census, 1975), and, more importantly, an increase in the extent to which women are gaining control over their lives, a trend which has broad psychosocial implications.

While fertility has been on the decline, female labor-force participation has increased. Today a majority of working-age women are employed outside the home. In fact, a majority of women with school-age children work, and about 42% of preschoolers' mothers work as well. Employment is even higher in single-parent families and among the never married (U.S. Department of Labor, 1977; U.S. Department of Commerce, 1979).

Of the numerous psychological changes we would expect to be associated with these new female role situations (Mason et al., 1976), one of the most theoretically interesting involves the well-known fact that women in our society consistently report higher levels of emotional distress than men in surveys of subjective well-being. As documented since World War II, women also have higher rates of clinically defined anxiety, some types of depression, and psychophysiological disorders (Weissman and Klerman, 1977; Goldman and Ravid, 1980; Link and Dohrenwend, 1980; Neugebauer et al., 1980). Since it is widely believed that these higher rates are linked to the fact that women find their positions in modern society more stressful and less rewarding than do men, we would expect this excess to decline to the extent that female roles are becoming more comparable to those of

* Address correspondence to Ronald C. Kessler at the Sociology Department or James A. McRae at the Institute for Social Research, The University of Michigan, Ann Arbor, MI 48106. Authors' names are listed alphabetically. We gratefully acknowledge the assistance of Roger Brown, Alexis Engel, and David M. Klingel and the comments of Walter Gove, Karen Mason, Peggy Thoits, Joseph Veroff, and the anonymous ASR reviewers. The data used here were generously provided by Elizabeth Douvan, Harold Dupuy, Richard Kulka, Ben Locke, Lenore Radloff and Joseph Veroff. Any errors of analysis or interpretation are, of course, ours and not those of these individuals. This work was partially supported by NIMH grants 5T32-MH14618 and 1R01-MH34479.

men (Gove, 1978; Makosky, 1980). Female attitudes toward childbearing and childrearing in relation to careers outside the home have also been changing. Dubnoff et al. (1976) report that employed women increasingly believe that their jobs do not detract from their effectiveness as mothers. The resolution of these sorts of potentially conflicting role pressures surely works to the mental health advantage of women.

To date, empirical work on the changing relationship between sex and mental illness has been scant. This is unfortunate because there is good reason to believe that the recent structural changes in women's roles have had mental health effects. Systematic study of these effects should shed light on the ways role strains lead to emotional distress, much as previous analyses have illuminated the changing relationships between gender and physical illnesses like peptic ulcer, diabetes, and coronary heart disease (Lewis and Lewis, 1977; Murphy, 1978).

Trend comparisons of male and female mental illness ought in particular to help disentangle the social-causation and selection interpretations that habitually confound cross-sectional comparisons. Female rates of mental illness are generally higher for those in traditional than nontraditional roles (Gove, 1978),[1] but it is not clear that role-related strains cause mental illness. It is equally possible that selection may be operating; that is, women in good mental health may simply be more likely to take up nontraditional roles than are women who are anxious or depressed.

Historic comparisons avoid this ambiguity by focussing on the effects of compositional shifts in role relationships. Here the causal influences of intra-individual selection processes pale, and observed changes in the relationship of sex to mental illness can unambiguously be interpreted as due to influences linked to roles.

EVIDENCE OF TRENDS

Since World War II, sociologists have documented a female preponderance in mental health treatment statistics, both among cases in community studies of true prevalence and in surveys of self-reported psychological distress. Gove and Tudor (1973) have argued that this preponderance was new with the postwar years and that women in earlier times had lower rates of "mental illness" than men. However, women do not have uniformly higher rates than men for all types of mental illness (see Neugebauer et al., 1980:54–5); and Gove and Tudor used a highly idiosyncratic definition of "mental illness," which combined types of disorders known to differ by sex with others known not to. The epidemiologic evidence on which Gove and Tudor based their claim has also been challenged, primarily on the basis of biases introduced by the use of key informants before the war and self-report measures afterwards (Dohrenwend and Dohrenwend, 1976, 1977). Although empirical work has been unable to document such biases (Clancy and Gove, 1974; Gove et al., 1976; Gove and Geerken, 1977b), the spotty evidence available suggests that a female preponderance in mood disorders was visible as far back as the nineteenth century (Weissman and Klerman, 1977:108).

We know that female rates of alcoholism, suicide, and property crime—behaviors that are traditionally much more common among men than women—have begun to rise since World War II (Weissman and Klerman, 1977; Gomberg, 1979; Sarri, 1979). These trends have been interpreted as evidence that women's roles are becoming more like those of men, but evidence about trends in the traditionally female disorders of anxiety, depression, and attempted suicide is much less adequate.

Weissman (1974:740) shows that the number of unsuccessful male suicide at-

[1] There is consistent evidence that children have a negative effect on the mental health of women (Pearlin, 1975; Radloff, 1975; Gove and Geerken, 1977a; Veroff, Douvan and Kulka, 1981). Employment outside the home seems to improve the mental health of women (Gove and Geerken, 1977a), although the evidence is mixed (Pearlin, 1975; Radloff, 1975; Newberry et al., 1979). Married women have relatively low rates when compared to the previously married. But they have only slightly lower rates than the never married (Fox, 1980; Veroff, Douvan and Kulka, 1981).

tempters has been on the rise in recent years and that the sex difference in these rates is declining, but this evidence is based on a small number of community studies and is not definitive. There is also evidence from the Midtown Manhattan restudy that the overall mental health of middle-aged women has improved relative to that of middle-aged men over the period 1954–74 (Srole and Fischer, 1980), but the panel on which this evidence is based is so severely censored that it is difficult to give it much weight.[2]

The only rigorous attempt to assess postwar trends in self-reported psychological distress is that of Veroff, Douvan, and Kulka (1981), who carried out a twenty-year update of the Americans View Their Mental Health study (Gurin et al., 1960) with a national probability survey administered in 1976. Focussing on problems in living, they found no evidence of a differential trend in the well-being of men and women. As we document below, though, when examined from a somewhat different perspective, their data do show that some aspects of women's well-being have improved relative to those of men.

DATA AND PLAN OF ANALYSIS

It is difficult to assess trends in overall mental health because good trend data are unavailable. Treatment statistics are risky to use, since sex differences in help-seeking confound estimates of relative male/female prevalence in these data (Kessler et al., 1981); in addition, differences in seeking help may vary over time. Epidemiologic surveys are preferable, but they do not contain information on a broad spectrum of mental health problems, only on symptoms of mild distress. There is good reason to believe, however, that changing role relationships will have a much more pronounced effect on the mild emotional problems measured in epidemiologic surveys than on other more serious types of disorders. As Gove (1978:190) suggests,

> . . . the neuroses and to a large extent the psychophysiological disorders are highly reactive to situational stress, whereas the psychoses, which tend to have a strong genetic component, are characteristically related (produced) by a very severe disruptive environment. As a consequence, we would expect the stress associated with the sex and marital roles of women to be much more strongly related to the neuroses and the psychophysiological disorders than to the psychotic disorders.

Therefore, in an effort to assess trends in the emotional distress of men and women since the Second World War, we turned to the epidemiological evidence. Two national surveys of mental health were conducted in the late 1950s, both of which are analyzed here. These are compared to a set of three surveys administered in the 1970s, of which one is a national study and two are community studies. Together, these five surveys span the period 1957 to 1976.

1. Americans View Their Mental Health (AVTMH), 1957. This survey was conducted by the Survey Research Center at the University of Michigan. The sample of 2440 was selected by means of a multistage cluster design and included noninstitutionalized adults in the United States aged 21 and over (see Gurin, Veroff, and Feld, 1960).

2. National Center for Health Statistics (NCHS), 1959–1961. This survey was carried out by NCHS as part of the ongoing Health Examination Survey. Sequential geographic cluster samples were drawn over a period of nearly three years. Together these samples are representative of the noninstitutionalized United States adult population. The sample of 7710 is limited to individuals aged 18–79 at the

[2] The 1954 Midtown sample was drawn in such a way as to be representative of the entire Midtown Manhattan population. The sample size was 1660. In 1974, Srole reinterviewed 695 of these respondents. In the original Midtown sample gender was "unrelated to mental morbidity" (Srole and Fischer, 1975:241). However, in the 695 reinterview respondents there was a strong relationship between symptom severity in 1954 and gender (Srole and Fischer, 1980:215). This discrepancy is probably due to the fact that married women who were initially interviewed when they were single are difficult to locate because of name changes. It would be possible to correct this bias partially by focussing only on respondents who were married in 1954, but this was not done in the Srole and Fischer analysis. We have been unable to obtain access to the data in order to carry out reanalysis.

time of the interview (see National Center for Health Statistics, 1970).

3 and 4. Community Mental Health Assessment (CMHA), 1971–1973. These two surveys were administered by the Center for Epidemiologic Studies, NIMH, in Kansas City, Missouri (KC) and Washington County, Maryland (WC). Sequential samples of adults aged 18 and over were drawn and interviewed from October, 1971 until the end of January, 1973 in Kansas City and from December, 1971 until the end of July, 1973 in Washington County. The sample sizes are 1173 in Kansas City and 1673 in Washington County (see Comstock and Helsing, 1973, 1976).

5. Americans View Their Mental Health Restudy (AVTMHR), 1976. This survey was designed as a 20-year cross-sectional replication of the original AVTMH survey. Like the baseline study in 1957, this replication was carried out by the Survey Research Center at Michigan. The sampling design is identical to the original. The sample size is 2230 (see Veroff, Douvan and Kulka, 1981).

Measures

All five surveys contain some indicators of the sort of emotional problems that are highly reactive to situational stress, mostly involving psychophysiological symptoms.[3] In NCHS, 10 self-report questions involving psychophysiological symptoms were asked; in the two AVTMH surveys, 19 of these questions were asked; and in the two CMHA surveys, the psychophysiological symptom questions contained in the Langner scale were asked. Some individual questions were identically worded in all of these surveys. However, the response categories varied from one study to the next, making it impossible to create identical scales or to compare trends in individual questions.

As a compromise, we created summed indexes of psychophysiological symptoms that are identical in the 1957 and 1976 AVTMH surveys and closely approximate the questions available in NCHS and CMHA.[4] Only the two AVTMH indexes can be compared rigorously. However, it is possible to use the other surveys to study changes in the relative levels of distress reported by men and women.

Analysis Plan

Our analysis consists of pooling data across surveys and estimating regression models in which time is a variable. Age of the respondent is also included to test the possibility that the relationship between sex and symptoms changed differentially across the age range. Stepwise procedures are used to arrive at a preferred model, one which cannot be significantly improved by including additional terms and which cannot be simplified without sig-

[3] Dohrenwend and Link (1980:116) show that the relationship between screening scales of this sort and clinical judgments is far from perfect. In a comparison of three studies containing both, "the proportion of cases identified by both procedures is less than half of all cases identified by one or both." However, even though this correspondence is imperfect, self-reported feelings of distress are important indicators of emotional difficulty in their own right.

[4] The items in the psychophysiological scales tap the following symptoms. AVTMH/AVTMHR: (1) ever have trouble getting to sleep or staying asleep; (2) ever been bothered by nervousness, feeling fidgety and tense; (3) ever troubled by headaches or pains in the head; (4) ever bothered by heart beating hard; (5) ever had spells of dizziness; (6) ever bothered by nightmares; (7) hands tremble enough to bother you; (8) troubled by hands sweaty so that you feel damp and clammy; (9) ever been times when you couldn't take care of things because you just couldn't get going.

NCHS: (1–9) same as in AVTMH/AVTMHR; (10) ever fainted or blacked out.

CMHA: (1–3) same as nos. 3, 4, and 7 in above scales; (4) feel weak all over much of the time; (5) suddenly feel hot all over; (6) ever bothered by cold sweats; (7) ever been times when personal worries get you down physically; (8) poor appetite; (9) fainting; (10) had fullness or clogging in head; (11) bothered by shortness of breath when not excited or working hard; (12) bothered by acid stomach several times a week.

The scale reliabilities, based on Heise and Bohrnstedt's (1970) Omega, are: (1) AVTMH: .795 (total pooled 1957–76 sample); .727 (males, 1957); .780 (females, 1957); .775 (males, 1976); .794 (females, 1976). (2) NCHS: .693 (total sample); .441 (males); .600 (females); (3) Kansas City: .599 (total sample); .551 (males); .670 (females); (4) Washington County: .631 (total sample); .709 (males); .855 (females).

nificantly worsening the fit. Models estimated include terms for time, sex, age in five year intervals (18–19, 20–24, . . ., 60–4, 65+), the square of age, and all possible two- and three-way interactions. Only hierarchical models are considered.

RESULTS

Table 1 presents the mean levels of psychophysiological symptoms (PPS) in the two AVTMH surveys. The indices have means of −.362 and .403 in 1957 and 1976, respectively. To evaluate trends, a series of OLS regression models was estimated. Table 2 shows R^2 coefficients for a few of these models.

We began by estimating a series of main-effects models. The most complete of these is Model 1. All of the parameters in this model are significantly associated with PPS. Next we added two-way interactions, one at a time. Only the interaction of decade and sex (X_1X_2) significantly improved on the main-effects model. This is shown as Model 2 in the table. Finally we looked at models containing three-way interactions—that is, interactions in which the joint effects of decade and sex are allowed to vary by age. None of these significantly improved on Model 2; Model 3, which is the most complete three-way interaction model, has an R^2 increment of only .001 over Model 2. Clearly, Model 2 is preferable. This model is

Table 2. Selected Models Considered in the Analysis of Psychophysiological Symptoms (PPS),

Model	Terms Included	R^2
(1)	X_1, X_2, X_3, X_3^2	.05878*
(2)	$X_1, X_2, X_3, X_3^2, X_1X_2$	.05998**
(3)	$X_1, X_2, X_3, X_3^2, X_1X_2, X_1X_3, X_2X_3, X_1X_3^2, X_2X_3^2, X_1X_2X_3, X_1X_2X_3^2$	.06132*

NOTE: X_1 denotes year (1 = 1957, 2 = 1976), X_2 denotes sex (0 = male, 1 = female), and X_3 denotes age (1 = 21–24, 2 = 25–29,...,10 = 65+).

* Significant at the .05 level.

** Significantly larger than R^2 for Model 1 at the .05 level.

$$\hat{Y} = \underset{}{1.021} + \underset{(.21)}{.373}\, X_1 + \underset{(.23)}{2.004}\, X_2 + \underset{(.12)}{.118}\, X_3 - \underset{(.01)}{.022}\, X_3^2 - \underset{(.32)}{.768}\, X_1X_2 \tag{1}$$

where $\hat{Y}$ is the expected value of PPS and the numbers in parentheses are the standard errors of the metric regression coefficients. (See the note to Table 2 for definitions of the X_i).

The predicted and observed mean levels of PPS are plotted in Figure 1. Women have higher levels of psychophysiological symptoms than men in both years. And this difference is constant over the cross-sectional age range within each year, with the average levels of PPS higher for older than younger people. Both men and women have higher PPS scores in 1976

Table 1. Mean Levels of Psychophysiological Symptoms (PPS) by Year, Age, and Sex

	Male				Female			
	1957		1976		1957		1976	
Age	$\bar{x}$	(n)	$\bar{x}$	(n)	$\bar{x}$	(n)	$\bar{x}$	(n)
21–24	−1.630	(65)	−.976	(102)	.328	(98)	.652	(140)
25–29	−2.383	(122)	−.760	(132)	.171	(168)	.631	(176)
30–34	−2.469	(130)	−1.629	(111)	.534	(175)	1.266	(149)
35–39	−1.543	(120)	−.931	(103)	.920	(158)	.973	(97)
40–44	−1.636	(121)	−1.945	(81)	.551	(149)	2.124	(93)
45–49	−1.819	(110)	−.286	(74)	.295	(134)	.480	(92)
50–54	−2.732	(98)	−1.276	(52)	.289	(114)	1.174	(101)
55–59	−1.287	(80)	−1.167	(81)	1.541	(96)	1.629	(100)
60–64	−2.629	(65)	.893	(74)	1.337	(87)	2.081	(90)
65+	−1.266	(159)	.130	(137)	2.300	(192)	1.820	(245)
TOTAL	−1.910	(1070)	−.773	(947)	.846	(1371)	1.280	(1283)

NOTE: The mean level of PPS across both years for men and women combined is 0.0. The standard deviations within years are 5.634 in 1957 and 5.309 in 1976. The standard deviation for both years combined is 5.494.

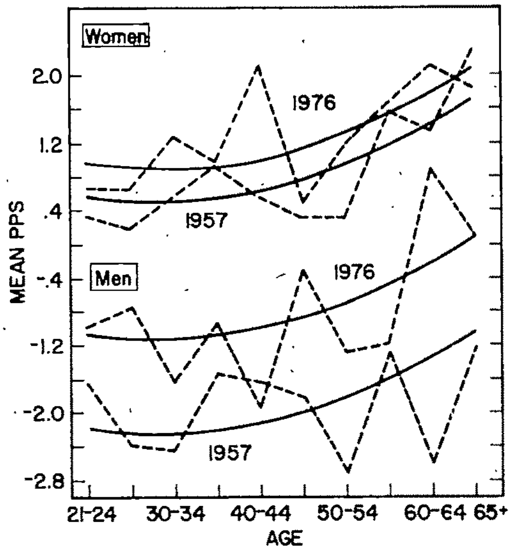


Figure 1. Expected and Observed Mean Levels of PPS by Year, Sex, and Age

than in 1957. However, the difference between the sexes is smaller in 1976 than in 1957. Moreover, this convergence is constant over the age range. In 1957, the mean female symptom level was 2.004 points higher than that of males, a difference close to a third of a standard deviation. By 1976, the difference had closed to 2.004 − .768 = 1.236 points, which is approximately two-tenths of a standard deviation. This represents a reduction of 38% in the 1957 gap.

For comparative purposes a series of age-by-sex models was estimated for NCHS and the two CMHA surveys. All three data sets, like AVTMH, show a sex effect that is constant over the cross-sectional age range. These effects, expressed in terms of standard deviation units, are presented in Table 3. It is quite clear that the difference between men and women was larger in the fifties than in the seventies. This apparent convergence might be due to an increase in the standard deviation of the symptom scale over time rather than to a reduction in the sex difference. However, as reported above, the standard deviation of PPS in the AVTMH surveys *decreased* slightly (from 5.634 in 1957 to 5.309 in 1976) over time. Therefore, the convergence we see in Table 3 probably underestimates slightly the convergence we would have been able to detect in a metric analysis of unstandardized symptom scales containing the same items across the total set of surveys considered.

It is important to note that we have not uncovered a pattern of improved mental health among women. Mean levels of PPS, as noted above, are higher in 1976 than in 1957. This relationship between year and symptoms reflects influences that transcend those which shape sex differences, since it is visible among women as well as men (for discussion of possible influences, see Kulka et al., 1979, and Veroff, Douvan, and Kulka, 1981). However, there has been an improvement in the relative PPS scores of women compared to men, an improvement that is visible not only in the AVTMH surveys but also in the three other studies reported in Table 3. It is this trend in relative symptom levels that is predicted by the sex role perspective. And it is this that we concentrate on in the following discussion.

Table 3. Standardized Differences in the Average Levels of Psychophysiological Distress Reported by Men and Women in Five Surveys

AVTMH 1957	NCHS 1959–1961	KC 1971–1973	WC 1971–1973	AVTMHR 1976
.36	.45	.27	.23	.22

NOTE: The AVTMH model is presented in the text as Equation 1. The preferred models for the other surveys are:

$$\text{(NCHS)}\ \hat{Y} = .279 + \underset{(.12)}{2.497}\,X_2 + \underset{(.08)}{.276}\,X_3 - \underset{(.01)}{.017}\,X_3^2;\ R^2 = .066$$

$$\text{(KC)}\ \hat{Y} = .008 + \underset{(.36)}{1.624}\,X_2 + \underset{(.05)}{.104}\,X_3;\ R^2 = .020$$

$$\text{(WC)}\ \hat{Y} = -.542 + \underset{(.28)}{1.265}\,X_2 + \underset{(.04)}{.163}\,X_3;\ R^2 = .021$$

Standard errors are in parentheses beneath the metric coefficients.

DISCUSSION

It is difficult to assess trends in overall mental health because good trend data are unavailable; in none of the surveys administered in the 1950s, for instance, did we find good screening scales of depression or psychological anxiety. On the basis of the available survey evidence, though, it is clear that the relationship between sex and psychophysiological symptoms has decreased over the past two decades. Evidence about suicides and suicide attempts also shows that men and women are becoming more alike in their rates.

We argued that trends of this sort could be predicted from compositional shifts in the roles of men and women that have taken place over the past two decades, especially trends toward greater female labor-force participation and toward lower rates of childbearing. These trends should decrease the strain-related types of mental health problems of women more than those of men. Although we have documented that rates of psychophysiological symptoms have become more comparable for men and women in the seventies than they were in the fifties, no evidence to support the compositional explanation has been presented.

We searched for such evidence by introducing explicit controls for composition into the regression model presented in Equation 1. Four sets of compositional variables were entered, first separately and then together. These are (a) five dummy variables to describe the interaction of sex with a three-category measure of marital status (married, previously married, and never married), (b) a series of variables to describe women who are the mothers of preschoolers or school-aged children,[5] (c) a series of variables to describe the educational attainments of men and women (years of education, a dummy for some high school, a dummy for high-school graduation, and a dummy for at least some college), and (d) a dummy variable to describe women who are employed outside their homes.

Only the dummy associated with female labor-force participation meaningfully reduced the interaction between decade and sex. In the revised version of Equation 1, female labor-force participation had a metric regression coefficient of −1.265 (s.e. = .217), which means that the typical employed woman has a PPS score 1.265 points lower than that of the typical homemaker. The interaction between decade and sex was reduced in this model from −.768 to −.599 (s.e. = .316), which means that the trend toward higher female labor-force participation over the two decades studied explains approximately 20% of the changing relationship between gender and psychophysiological symptoms.[6]

These influences might be partly due to shifts in the educational attainments of women over the two decades. But including years of education—dummy variables for having 9–11, 12, or 13–15 years of schooling—and the interactions of these variables with sex in the model reduces the effect of female labor-force participation only slightly—from −1.265 to −1.096.

A somewhat more complicated model was tested to see if the relationship between female labor-force participation and PPS had changed over time. This is a particularly important possibility to evaluate because it helps disentangle the role-strain and selection interpretations mentioned at the beginning. If it is true that women who are mentally healthy are more likely to

[5] We took an empirical approach to conceptualizing the influence of children. From the data available in both the 1957 and 1976 surveys we were able to create the following measures: (a) total number of children had (not necessarily living at home); (b) a dummy for at least one preschooler at home; (c) a dummy for at least one child 5–16 at home; (d) a dummy for only preschooler(s) at home; (e) a dummy for only child(ren) 5–16 at home; (f) a dummy for only child(ren) 17+ at home; (g) a dummy for only child(ren) 0–16 at home; (h) a dummy for 3 or more children, all of whom are preschoolers, at home; (i) a dummy for 3 or more children, all of whom are 5–16, at home; (j) a dummy for 3 or more children, all of whom are 0–16, at home. Unlike the other control variables (for marital status, employment, education), which were entered listwise in the regression equation, these controls were entered stepwise.

[6] The coefficient −.768 comes from Equation 1. The reduction from −.768 to −.599 is one of .169, which is 22% of .768.

enter the labor force than other women, we would expect that the relationship between labor-force participation and symptoms of PPS would be lower in the seventies than in the fifties, since selection processes of this sort will necessarily attenuate as female labor-force participation becomes more common.[7]

The data show that the 1.265 difference between the symptoms of employed women and homemakers is constant across the two decades, a result which clearly refutes the argument that selection alone is at work. This result is consistent with the view that employment outside the home provides women with a variety of concrete and symbolic rewards unavailable to homemakers and that it allows them to escape from a variety of strains to which homemakers are exposed. The constant difference is also consistent with the prediction that as female labor-force participation increases the mental health advantages of employment will not diminish.

The other compositional changes—in rates of intact marriage, childbearing, and educational attainment—had comparatively trivial effects on the relationship between gender and psychophysiological symptoms. This means that together the compositional variables we considered account for only about 20% of the convergence documented above. Although a more refined compositional analysis might explain a larger percentage of the trend toward sex parity in PPS, we suspect that compositional shifts of this sort are not primarily responsible for the convergence. The gap is closing equally for the married, never married, and previously married, for those who have children and those who do not, for the young and the old. It is unlikely that a pervasive trend of this sort represents the influence of changes in the different role relationships between men and women that exist across all these varied sectors of society.

One possibility is that the role changes that women have experienced over the past two decades are helping to close the male-female mental health gap largely by increasing the distress of men rather than decreasing it for women. This possibility would help account for why the specifications we considered—all of which focus on women—play such a little part in the year-by-sex interaction. It is consistent with the results shown in Figure 1, where we see that mean levels of PPS increased a great deal for men between 1957 and 1976 and only to a lesser extent for women. It is also consistent with Rosenfield's (1980) recent finding that the average depression levels reported by husbands of employed women are higher than those reported by the husbands of fulltime homemakers. However, the evidence on this point is currently scant.

Another plausible explanation for the trend toward parity in PPS is that, as the role possibilities for women expand, women in many role situations increasingly find it possible to adopt the type of life they would like to have. This explanation is consistent with evidence about the mental health advantages of "fit" between a person and his or her environment (French et al., 1974) and with evidence that women today are more likely than those in the past to report that they prefer the roles they currently occupy to other roles they might have had (Dubnoff et al., 1978).

The importance of the expanding set of role possibilities, the changing role relationships between men and women, and changing attitudes among and about women in bringing about the pervasive trend toward parity in psychophysiological symptoms cannot be assessed with the data we have available to us. Further monitoring of the trend will be required before we can deal with these issues adequately. However, even though the driv-

[7] If PPS influences employment status (ES) but ES does not influence PPS, the slope of ES on PPS will be constant across year, and the slope of PPS on ES will be lower in 1976 than in 1957. The second result holds because the variance of PPS will not increase in proportion to changes in ES. On the other hand, if ES influences PPS but PPS does not influence ES, the slope of PPS on ES will be constant across year. If both effects are present (ES influences PPS and PPS influences ES) changes in the slopes can take on any value.

Since we find that the slope of PPS on ES remains constant across year, our results imply that we can exclude the possibility that selection alone (that is, PPS influencing ES but ES not influencing PPS) is at work.

ing forces behind the trend remain unclear, the trend itself is now clearly documented.

REFERENCES

Clancy, Kevin and Walter R. Gove
1974 "Sex differences in respondents' reports of psychiatric symptoms: an analysis of response bias." American Sociological Review 80:205–16.

Comstock, G. W. and K. Helsing
1973 "Characteristics of respondents and nonrespondents to a questionnaire for estimating community mood." American Journal of Epidemiology 97:233.
1976 "Symptoms of depression in two communities." Psychological Medicine 6: 551–63.

David, H. P. and W. P. Baldwin
1979 "Childbearing and child development: demographic and psychosocial trends." American Psychologist 34:866–71.

Dohrenwend, B. P. and B. S. Dohrenwend
1976 "Sex differences and psychiatric disorders." American Sociological Review 81:1447–54.
1977 "Reply to Gove and Tudor's comment on 'Sex differences and psychiatric disorders.'" American Sociological Review 82:1336–45.

Dubnoff, S. J., J. Veroff, and R. A. Kulka
1978 "Adjustment to work: 1957–76." Paper presented at the meeting of the American Psychological Association, Toronto.

Engel, A., R. S. Murphy, K. Maurer, and E. Collins
1978 Plan and operation of the HANES I augmentation survey of adults 25–74 years: United States, 1974–75. Vital and Health Statistics, Programs and Collection Procedures. Washington, D.C.: Series 1, No. 14.

Fox, J. W.
1980 "Gove's specific sex-role theory of mental illness: a research note." Journal of Health and Social Behavior 21:260–7.

French, J. R. P., Jr., W. Rodgers, and S. Cobb
1974 "Adjustment as person-environment fit." Pp. 316–33 in George V. Coelho, David A. Hamburg, and John E. Adams (eds.), Coping and Adaptation. New York: Basic Books.

Goldman, N. and R. Ravid
1980 "Community surveys: sex differences in mental illness." Pp. 31–56 in Marcia Guttentag, Susan Salasin and Deborah Belle (eds.), The Mental Health of Women. New York: Academic Press.

Gomberg, E. S.
1979 "Problems with alcohol and other drugs." Pp. 204–40 in Edith S. Gomberg and Violet Franks (eds.), Gender and Disordered Behavior. New York: Brunner/Mazel.

Gove, W. R.
1978 "Sex differences in mental illness among adult men and women: an evaluation of four questions raised regarding the evidence on the higher rates of women." Social Science and Medicine 12B:187–98.

Gove, W. R. and M. R. Geerken
1977a "The effect of children and employment on the mental health of married men and women." Social Forces 56:66–76.
1977b "Response bias in surveys of mental health: an empirical investigation." American Journal of Sociology 82:1289–317.

Gove, W. R., J. McCorkel, T. Fain, and M. Hughes
1976 "Response bias in community studies of mental health: systematic bias or random noise?" Social Science and Medicine 13: 55–67.

Gove, W. R. and J. F. Tudor
1973 "Adult sex roles and mental illness." American Journal of Sociology 98:812–35.

Gurin G., J. Veroff, and S. Feld
1960 Americans View Their Mental Health. New York: Basic Books.

Heise, D. R. and G. W. Bohrnstedt
1970 "Validity, invalidity, and reliability." Pp. 104–29 in Edgar F. Borgatta and George W. Bohrnstedt (eds.), Sociological Methodology 1970. San Francisco: Jossey-Bass.

Kessler, R. C., R. Brown, and C. Broman
1981 "Sex differences in psychiatric help-seeking: evidence from four large-scale surveys." Journal of Health and Social Behavior 22:49–64.

Kulka, R. A., J. Veroff, and E. Douvan
1979 "Social class and the use of professional help for personal problems: 1957 and 1976." Journal of Health and Social Behavior 20: 2–16.

Lewis, C. E. and M. A. Lewis
1977 "The potential impact of sexual equality on health." The New England Journal of Medicine 297:863–9.

Link, B. and B. P. Dohrenwend
1980 "Formulation of hypotheses about the true prevalence of demoralization in the United States." Pp. 114–32 in Bruce P. Dohrenwend, Barbara S. Dohrenwend, Madelyn S. Gould, Bruce Link, Richard Neugebauer, and Robin Wunsch-Hitzig (eds.), Mental Illness in the United States: Epidemiological Estimates. New York: Praeger.

Makosky, V. P.
1980 "Stress and the mental health of women: a discussion of research and issues." Pp. 111–28 in Marcia Guttentag, Susan Salasin and Deborah Belle (eds.), The Mental Health of Women. New York: Academic Press.

Mason, K. O., J. L. Czajka, and S. Arber
1976 "Change in U.S. women's sex-role attitudes, 1967–1974." American Sociological Review 41:573–96.

Murphy, H. B. M.
1978 "Changing sex ratios in illness." Social Science and Medicine 12:143–50.

National Center for Health Statistics
1951 Vital Statistics for the United States (Vol. 1). Washington, D.C.: U.S. Department of Health, Education, and Welfare.

1969 Vital Statistics for the United States (Vol. 1) Washington, D.C.: U. S. Department of Health, Education, and Welfare.

1970 Selected Symptoms of Psychological Distress. Public Health Service Publication No. 1000, Series II, No. 37. Washington, D. C.: U.S. Government Printing Office.

1979 "Final natality statistics, 1977." Monthly Vital Statistics Report 27(12). (Supplement).

Neugebauer, R., B. P. Dohrenwend, and B. S. Dohrenwend
1980 "Formulation of hypotheses about the true prevalence of functional psychiatric disorders among adults in the United States." Pp. 45–94 in Bruce P. Dohrenwend, Barbara S. Dohrenwend, Madelyn S. Gould, Bruce Link, Richard Neugebauer and Robin Wunsch-Hitzig (eds.), Mental Illness in the United States: Epidemiological Estimates. New York: Praeger.

Newberry, P., M. M. Weissman, and J. K. Myers
1979 "Working wives and housewives: do they differ in mental states and social adjustment." Draft, Department of Psychiatry, Yale University.

Pearlin, L. I.
1975 "Sex roles and depression." Pp. 191–207 in N. Datan and L. Ginsberg (eds.), Proceedings of Fourth Life-Span Developmental Psychology Conference: Normative Life Crises. New York: Academic Press.

Radloff, L. S.
1975 "Sex differences in depression: the effects of occupation and marital status." Sex Roles 1:249–65.

Rosenfield, S.
1980 "Sex differences in depression: do women always have higher rates?" Journal of Health and Social Behavior 21:33–42.

Sarri, R. C.
1979 "Crime and the female offender." Pp. 159–203 in Edith S. Gomberg and Violet Franks (eds.), Gender and Disordered Behavior. New York: Brunner/Mazel.

Srole, L. and A. K. Fischer (eds.)
1975 Mental Health in the Metropolis: The Midtown Manhattan Study. Revised and Enlarged Edition. New York: Harper and Row.

Srole, L. and A. K. Fischer
1980 "The 'Midtown Manhattan longitudinal study' vs 'The mental paradise lost doctrine.'" Archives of General Psychiatry 37:209–21.

U.S. Bureau of the Census
1975 "Fertility expectations of american women: June 1974." Current Population Reports, Series P-20, No. 277. Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Commerce, Bureau of the Census
1979 Population Profile of the United States: 1978; population characteristics (Current Population Reports, Series P-20, No. 336). Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Labor, Women's Bureau
1977 Working Mothers and their Children. Washington, D.C.: U.S. Government Printing Office.

Veroff, J., E. Douvan, and R. Kulka
1981 The Inner American. New York: Basic Books.

Weissman, M. M.
1974 "The epidemiology of suicide attempts, 1960 to 1971." Archives of General Psychiatry 30:727–46.

Weissman, J. M. and G. L. Klerman
1977 "Epidemiology of mental disorders." Archives of General Psychiatry 35:705–12.

RESEARCH NOTES

TRENDS IN UNITED STATES MEN'S AND WOMEN'S SEX-ROLE ATTITUDES: 1972 to 1978*

ANDREW CHERLIN
PAMELA BARNHOUSE WALTERS
Johns Hopkins University

Women's sex-role attitudes have shown steadily increasing support for nontraditional sex roles from the early 1960s to early 1970s, and their sex-role attitudes have become increasingly consistent. Data in this study show that these trends for white women have continued from 1972 to 1975 and that the levels of egalitarianism and attitude consistency attained in 1975 were maintained through 1978. White men exhibited as much change in the egalitarian direction as did white women between 1972 and 1978; by 1978, white men's sex-role attitudes resembled those of white women. White men's sex-role attitudes also became more consistent between 1972 and 1975. Changes in sex-role attitude consistency for whites could not be accounted for by changes in liberal-conservative political attitudes. Black men's and women's sex-role attitudes showed no significant changes between 1972 and 1978; however, small sample size precluded detailed analyses of these trends.

We know from several studies that American women's attitudes concerning appropriate sex roles and women's rights have become less traditional since the early 1960s. For example, Mason, Czajka, and Arber's (1976) analysis of responses of women to questions on women's roles in five surveys undertaken between 1964 and 1974 documented a consistent trend toward more egalitarian attitudes. Yet the years between the mid-1960s and the early 1970s also were a time of protests against the Vietnam War, of an economic boom fueled by the war, of increased civil rights activity, and of a revival of feminism. Since then, however, the war has ended, the economy has not performed as well, civil rights activity appears to have slowed, and conservative protests against the Equal Rights Amendment, abortion, and other feminist issues have increased. Conventional wisdom suggests that, correspondingly, social and political attitudes have become more conservative since the early 1970s, and some survey evidence supports this belief (see, for example, Davis, 1980). In this paper, we will investigate whether the trend of the 1960s and early 1970s toward more egalitarian sex-role attitudes continued into the late 1970s despite a general conservative drift in social and political sentiment. Furthermore, we will address variations in the trend according to race and sex, and we will look at trends in the consistency of sex-role attitudes.

DATA

Our source of data is the NORC General Social Survey, an annual personal interview sampling of the United States adult population from 1972 to 1978. The sampling universe of the GSS is the total noninstitutionalized English-speaking population of the United States, 18 years of age or older. A different sample of individuals

* Direct all correspondence to Andrew Cherlin, Department of Social Relations, Johns Hopkins University, Baltimore, MD 21218.

This material is based upon work supported in part by the Social Science Research Council and by the National Science Foundation under Grant No. 77-21686. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the Council or the Foundation. We also wish to thank Glenna Spitze and Arland Thornton for helpful comments on earlier versions.

was selected every year. Sample sizes averaged 1,520, with a total of 10,652 cases from 1972 to 1978.

In the 1972, 1974, 1975, 1977, and 1978 interviews, the GSS included two standard survey items which have become the most widely cited indicators of changes in sex-role attitudes. These are the only questions relating to sex-roles which have been asked in national surveys with any regularity over the past few decades. The first of these questions is "If your party nominated a woman for President, would you vote for her if she were qualified for the job?" (We will sometimes refer to this question by its GSS mnemonic FEPRES.) As others have noted (Duncan and Duncan, 1978; Spitze and Huber, 1980), the proportion saying they would vote for a woman for president increased gradually between 1937 and the late 1950s. It stayed roughly constant between the late 1950s and the late 1960s; then it increased sharply between 1969 and 1974; and since 1974 it has been increasing slowly again.

The second question (FEWORK) is "Do you approve or disapprove of a married woman earning money in business or industry if she has a husband capable of supporting her?" This question was asked of national samples in 1938, 1945, 1970, and in the GSS. The proportion of the population approving of a married woman earning money in business or industry appears to have increased linearly; a weighted regression of the proportions on the survey year yields an R^2 of .98 (Smith, 1980).

In addition, the GSS included in 1974, 1975, 1977, and 1978 two further questions on women and politics. They were "Do you agree or disagree with this statement? Women should take care of running their homes and leave running the country up to men" (FEHOME), and "Tell me if you agree or disagree with this statement: Most men are better suited emotionally for politics than are most women" (FEPOL). We factor analyzed all four sex-role items using a principal factoring method followed by a varimax rotation, and we found that the three political items (FEPRES, FEHOME, and FEPOL) loaded on one factor and the work item (FEWORK) loaded on a second factor. As Spitze and Huber (1980) have argued, FEPRES and FEWORK appear to represent different dimensions of sex-role attitudes. FEPRES and the two other political items, FEPOL and FEHOME, refer to women's political rights, and they may be indicators of prejudice (see also Ferree, 1974). On the other hand, FEWORK relates to economic concerns and to the fear that women may take jobs away from men. In our analyses, we will consider FEWORK and the political items as representing two different dimensions of sex-role attitudes, and we will report trends in the consistency of men's and women's responses on these two dimensions.[1] Unlike some previous studies (Mason et al., 1976; Thornton and Freedman, 1979b), we were unable to include a dimension measuring attitudes toward women's roles at home. Yet the items in the GSS are similar to those employed by many previous researchers, and in some cases they are identical. We would expect, then, that the trends we report here would not change significantly if alternative indicators of the same sex-role attitude dimensions were employed.

Our analyses of attitude consistency also make use of two indicators of trends in general political attitudes. The first is a six-item scale composed of the responses in 1974, 1975, and 1978 to a set of questions on national spending priorities. The GSS respondents were asked whether the country is spending too much, too little, or about the right amount on programs such as "solving the problems of the big cities," "improving the conditions of blacks," and so forth. There were eleven such items in the GSS; we chose six which differentiated most clearly between liberal and conservative positions.[2] The second

[1] We scored all four items such that a traditional response was coded two and a nontraditional response was coded one. We then summed the scores on FEPRES, FEHOME, and FEPOL to form a three-item political sex-role scale. (Persons who answered "don't know" on any of the items were omitted.) The scale had a reliability for white men of .68 and .64 for 1974 and 1978 and .71 for white women in both survey years. Results of our factor analysis (and all analyses referenced in subsequent footnotes) are available from the authors on request.

[2] We factor analyzed the eight spending priority items that, according to Davis (1980), differentiate

Table 1. Change in the Proportion Approving Nontraditional Sex-Roles, by Race and Sex, 1972 and 1978

A. Proportion Approving of a Married Woman Earning Money in Business or Industry if She Has a Husband capable of Supporting Her

	1972	1974	1975	1977	1978	Change 1972–1975	Change 1975–1978	Change 1972–1978
White men	.62	.66	.70	.68	.73	.09**	.02	.11**
	(662)	(606)	(588)	(605)	(560)			
Black men	.68	.62	.62	.63	.64	−.06	.02	−.04
	(128)	(71)	(69)	(67)	(64)			
White women	.67	.73	.72	.66	.75	.06**	.02	.08**
	(657)	(669)	(711)	(713)	(778)			
Black women	.76	.65	.72	.64	.75	−.04	.02	−.02
	(127)	(96)	(90)	(106)	(91)			

B. Proportion Who Would Vote For a Woman For President if She Were Qualified and Nominated by Their Party

	1972	1974	1975	1977	1978	Change 1972–1975	Change 1975–1978	Change 1972–1978
White men	.73	.82	.83	.83	.83	.10**	.00	.10**
	(638)	(601)	(579)	(597)	(559)			
Black men	.77	.71	.77	.75	.82	.00	.05	.05
	(121)	(70)	(70)	(68)	(62)			
White women	.72	.80	.79	.77	.79	.07**	.00	.07**
	(646)	(657)	(698)	(699)	(768)			
Black women	.83	.83	.78	.80	.90	−.06	.12*	.07
	(124)	(98)	(89)	(105)	(88)			

NOTE: Proportion reported is the number who responded "yes" divided by the number who responded "yes" or "no". Further investigation showed that the pattern of attitude change was very similar when "don't know" responses were included in the denominator. Ns in parentheses.

* $p < .05$ based on t-test of differences between proportions

** $p < .01$ based on t-test of differences between proportions

indicator was a seven-point scale on which respondents were asked to rate themselves from extremely liberal—point one—to extremely conservative—point seven.

between a liberal and conservative position. A large drop in eigenvalues between the first and second factor clearly suggested a one-factor solution, and the first factor contained moderate to high loadings for seven items. Further analysis showed that dropping one of these seven items improved the reliability of the resulting scale. We formed a single spending-priorities scale by rescoring all items such that a high score indicated a conservative attitude and summing across all six responses. Scale reliabilities were .61 and .55 for white men and women, respectively, in 1974, and .63 and .54 for white men and women in 1978. All the items were prefaced with the following statement: "We are faced with many problems in this country, none of which can be solved easily or inexpensively. I'm going to name some of these problems, and for each one I'd like you to tell me whether you think we're spending too much money on it, too little money, or about the right amount." The six items included in our scale are improving and protecting the environment, improving and protecting the nation's health, solving the problems of the big cities, improving the nation's education system, improving the conditions of blacks, and welfare.

RESULTS

Attitude Change from 1972 to 1978. In Table 1 we present the proportions giving nontraditional responses to FEWORK and FEPRES, from 1972 to 1978, by sex and race. We can see that between 1972 and 1978 there was substantial movement toward an egalitarian position among whites ($p<.01$ for men and women for both items). Almost all of the movement occurred between 1972 and 1975; from 1975 to 1978, there were no significant changes. Among blacks, however, 1978 responses were not significantly different from the 1972 responses. We would suggest that one possible reason for the apparently greater change among whites than among blacks might be that during this period the labor-force participation rate of white women was increasing steadily while the rate for black women was relatively stable (U.S. President, 1979). We know very little about trends in sex-role attitudes for blacks, because data limitations have forced many scholars to restrict their attention to whites (Mason et

al., 1976; Thornton and Freedman, 1979a, 1979b).

Looking at the whites, we see that, although both men and women became significantly less traditional on both of these attitude dimensions between 1972 and 1978, attitude change appears to have been somewhat greater among men. In 1972, white men were less likely than white women to approve of a married woman working ($p<.05$), but by 1978 there was no statistically significant difference between the sexes on this item. White men and women were about equally likely to endorse voting for a woman for president in 1972, but by 1978 white men were less traditional than white women ($p < .10$). Yet we know that between 1969 and 1972, the shift in responses to FEPRES was greater for women than for men. (The proportion saying they would vote for a woman increased by .19 for women and .11 for men between 1969 and 1972 (Spitze and Huber, 1980).) Thus, according to the responses to these two questions, white men seem to have changed their sex-role attitudes as much as or perhaps more than white women during the mid-1970s, reversing the pattern of the late 1960s and early 1970s.

We investigated this apparent sex difference among whites further with a multivariate model which also included age and educational attainment. A series of linear models of the log-odds of responding nontraditionally to FEWORK and FEPRES were estimated using sex, age, education, and year as independent variables (see Table 2). The independent variables were coded as follows: for sex, 0 = male and 1 = female; for year, 0 = 1972 and 1 = 1978. Education was defined as years of school completed and coded such that 0 = 12 years or less and 1 = more than 12 years. Age was trichotomized and required two dummy vectors for coding: 18–30 year olds were coded as 0 on both vectors, 31–50 year olds were coded as 1 on the first vector, and those over 50 years old were coded as 1 on the second vector. Interaction terms were formed by multiplicative combinations of the appropriate main-effect dummy vectors. Maximum likelihood estimates for these models,

Table 2. Estimated Logit Coefficients

	Dependent Variables			
	FEWORK		FEPRES	
Independent Variables	Coefficient	T-Statistic	Coefficient	T-Statistic
Main Effects				
(S) Sex (female)	.42*	2.04	.16	.64
(Y) Year (1978)	.44*	2.97	.68*	2.16
(E) Education (more than 12 years)	1.12**	6.88	.22	1.43
Age				
(A1) (31–50 years old)	−.13	−.76	−.58**	−2.52
(A2) (over 50 years old)	−.60**	−3.67	−.21	−.89
Interaction Effects				
SY	.07	.34	.32	.82
SE	.05	.14	—	—
SA1	.30	1.27	−.01	−.03
SA2	−.36	−1.56	−.66*	−2.51
YE	.34	1.16	.22	.97
YA1	—	—	.20	.52
YA2	—	—	−.72**	−2.60
SYE	−.87	−1.83	—	—
SYA1	—	—	−.50	−.97
SYA2	—	—	−.35	−1.44
Constant	.37		1.30	
χ^2	9.40		9.47	
df	12		10	
	p=.67		p=.49	

NOTE: FEWORK and FEPRES are coded as follows: 1 = nontraditional response; 0 = traditional response. The mean proportions of nontraditional responses are FEWORK, .69, and FEPRES, .77.

* $p < .05$

** $p < .01$

which are usually called logit models, were obtained by the Newton-Raphson algorithm (Haberman, 1978). For FEWORK and FEPRES, we selected the most parsimonious models that contained all of the terms necessary for a satisfactory fit.

For FEPRES, the best-fitting model included interaction terms involving sex, year, and age; these interaction terms could not be deleted without significantly worsening the fit of the model. They indicated that the sex difference in changes on FEPRES was most pronounced for men and women in their thirties and forties. This sex difference in the amount of change, however, was small in substantive terms. For example, under the model in Table 2, the predicted probabilities of giving the egalitarian response to FEPRES increased between 1972 and 1978 by .16 for white men aged 31 to 50 and by .13 for white women aged 31 to 50. In the younger and older age groups, the predicted probabilities increased by virtually the same amount for men and for women.

For FEWORK, the best-fitting model included interaction terms involving sex, year, and education; these interaction terms implied that the difference between the sexes in the amount of change on FEWORK occurred only among those with at least one year of college. Because of the presence of numerous interaction terms, however, it is difficult to interpret the coefficients in Table 2 directly. The rather unexpected pattern of change in FEWORK according to sex and education may be more clear if, informed by the logit model, we return to cross-tabulations of the GSS data. Table 3 shows that between 1972 and 1978 men at both educational levels and women with 12 years of schooling or less became significantly less traditional in their attitudes toward women working ($p < .05$ in each case), but women with at least one year of college became slightly more traditional—particularly between 1975 and 1978. We do not know why college-educated women, the group that most observers would expect to be in the vanguard of change, appear to have showed less movement toward a nontraditional position. Overall, then, our more detailed analysis indicates that the only substantial difference between men and women in the amount of change between 1972 and 1978 occurred in the responses of those with one or more years of college to the FEWORK item.

Attitude Consistency. If each person's sex-role attitudes were completely consistent, then we would expect to find a very strong association between nontraditional responses to FEWORK and nontraditional responses to FEPRES. Previous studies have reported evidence from a series of cross-sectional surveys (Mason et al., 1976) and from panel studies (Thornton and Freedman, 1979b) that white women's sex-role attitudes became more consistent during the 1960s and early 1970s. In Table 4 we present a measure of association, "gamma," between responses to FEPRES and FEWORK by race and sex for 1972, 1974, 1975, 1977, and 1978.

Table 3. Proportion of Whites Approving of a Married Woman Earning Money, by Sex and Education, 1972 to 1978

	1972	1975	1978	Change 1972–1975	Change 1975–1978	Change 1972–1978
Men, 12 years or less	.53 (421)	.60 (369)	.63 (337)	.07*	.03	.10**
Men, 13 or more years	.78 (240)	.88 (218)	.89 (221)	.10**	.01	.11**
Women, 12 years or less	.61 (492)	.65 (510)	.71 (552)	.04	.06*	.10**
Women, 13 or more years	.85 (165)	.91 (199)	.83 (223)	.06	−.08*	−.02

NOTE: Proportion reported is the number who responded "approve" divided by the number who responded "approve" or "disapprove." Ns in parentheses.
* $p < .05$ based on t-test of differences between proportions
** $p < .01$ based on t-test of differences between proportions

Table 4. Gammas between Responses to FEWORK and Responses to FEPRES, by Race and Sex, 1972 to 1978

	1972	1974	1975	1977	1978
White men	.16	.40	.56	.39	.57
	(627)	(588)	(571)	(588)	(550)
Black men	.31	.37	.29	.62	.49
	(121)	(69)	(69)	(66)	(68)
White women	.24	.45	.62	.48	.65
	(629)	(644)	(684)	(690)	(757)
Black women	.38	.22	.54	.31	.13
	(122)	(101)	(91)	(104)	(95)

NOTE: Subjects responding "don't know" or not responding were omitted from these calculations. A high gamma indicates that a respondent who gave a nontraditional (or traditional) reply to FEWORK was likely to give a nontraditional (or traditional) reply to FEPRES. Ns in parentheses.

"Gamma" is a simple transformation of the odds ratio for the two-by-two table of responses to FEPRES by responses to FEWORK. (We eliminated cases in which the respondents answered "don't know.") Duncan and Duncan (1978) present a similar table of odds ratios for men and women (not broken down by race) for 1972, 1974, and 1975. As Table 4 demonstrates, the strength of the relationship increased sharply for white men and women between 1972 and 1975, but the increase for whites between 1975 and 1978 was negligible. (The temporary decline in the consistency of responses to FEWORK and FEPRES among whites in 1977 appears to be the result of a sharper decline that year in egalitarian responses to FEWORK than to FEPRES.) In every year, the gammas were higher for white women than for white men, suggesting that white women tended to have more consistent attitudes along these two dimensions than did white men. For blacks, however, the gammas fluctuated greatly, and no clear trend in attitude consistency is apparent. Thus, we would conclude from Table 4 that there was an increase in sex-role attitude consistency for white women and white men between 1972 and 1975.

For the 1974 to 1978 period, we can examine the consistency among a larger set of items: FEWORK, the three-item scale of political sex-role attitudes (which includes FEPRES, FEPOL, and FEHOME), the six-item scale of national spending priorities (which we take to be a general measure of conservative versus liberal attitudes), and the self-identification on a seven-point liberal-conservative scale. We must limit this part of our analysis to whites, however, because FEPOL was only asked of half the sample in 1974, thus further reducing the already limited amount of data on blacks. In Table 5 we present the correlations

Table 5. Product-Moment Correlations among Indicators of Sex Roles, Spending Priorities, and Political Self-Identification, for White Men and Women, 1974, 1975, and 1978

White Women				
Correlations	1974	1975	1978	Change 1974–1978
FEWORK/Political sex roles	.20	.39	.38	.17
FEWORK/Spending priorities	.07	.10	.11	.04
FEWORK/Political views	.09	.14	.05	−.04
Political sex roles/Spending priorities	.17	.23	.16	−.01
Political sex roles/Political views	.16	.13	.13	−.02
Spending priorities/Political views	.27	.23	.23	−.04
	(n = 209)	(n = 429)	(n = 505)	
White Men				
Correlations	1974	1975	1978	1974–1978
FEWORK/Political sex roles	.30	.31	.30	−.00
FEWORK/Spending priorities	.16	.01	.02	−.14
FEWORK/Political views	.17	.08	.14	−.03
Political sex roles/Spending priorities	.22	.18	.08	−.14
Political sex roles/Political views	.29	.12	.09	−.20
Spending priorities/Political views	.33	.32	.32	−.01
	(n = 216)	(n = 405)	(n = 401)	

NOTE: "Political sex roles" refers to the 3-item political sex roles scale. "Spending priorities" is the 6-item scale of responses to questions on federal spending priorities. "Political views" refers to respondents' self-identification along a liberal/conservative political dimension. All items and scales were recoded such that a high score indicates a traditional attitude.

among these four variables for white women and men in 1974, 1975, and 1978.[3] Looking first at white women, we see that the correlation between the two sex-role variables increased from .20 to .38 during the four-year interval, confirming again the trend toward greater consistency in sex-role attitudes. Yet, on closer examination, we see that all of the increase in consistency had taken place by the 1975 interview. After 1975, the level of consistency remained constant. If the rise in sex-role attitude consistency was simply a function of a more general trend toward greater consistency of social and political attitudes, we would expect to see an increased association between sex-role attitudes and general political views during the 1970s. Most of the correlations, however, between the two sex-role variables, on the one hand, and the two variables measuring general conservative versus liberal preferences, on the other hand, decreased slightly between 1974 and 1978. The GSS data, consequently, suggest that the trend toward greater sex-role attitude consistency among white women was not merely a reflection of a more general increase in political attitude consistency. Rather, the data suggest that white women's sex-role attitudes were becoming a distinct and increasingly coherent body of opinion. For white men, however, there was essentially no change in the correlation between the two sex-role variables between 1974 and 1978. During that time, the correlations between the sex-role variables, on the one hand, and the more general conservative versus liberal measures, on the other hand, declined for men. Table 5, then, shows no evidence of increased sex-role attitude consistency for white men between 1974 and 1978—although Table 4 suggests some increase in consistency between 1972 and 1975.

[3] Following Nie *et al.* (1976) and Thornton and Freedman (1979b), we consider the correlation among attitude items to be a measure of attitude constraint. Barton and Parsons (1977), however, argued that the inter-item correlation is an inappropriate measure of attitude constraint if the subgroups differ in degree of attitude heterogeneity. Differences in attitude heterogeneity were relatively small for the subgroups considered here; for each of the variables in Table 5, the variances were similar across categories of year and sex, thus rendering the inter-item (product-moment) correlation an appropriate measure of attitude constraint.

SUMMARY AND DISCUSSION

The GSS data imply that the sex-role attitudes of white men and white women were less traditional, on average, in 1978 than in 1972. Almost all of the change, however, occurred between 1972 and 1975; after 1975 attitudes remained essentially constant. By 1978 there was little difference between the proportions of white men and white women approving of a married woman earning money in business or industry if she had a husband capable of supporting her, and white men were slightly more likely to say that they would vote for a woman for president than were white women. Furthermore, white men with at least some college education showed more change in the proportion approving of a married woman working than did similarly-educated white women.

White men's and women's sex-role attitudes also appear to have become more consistent during the period of the GSS surveys, especially between 1972 and 1975. It appears that the trend toward greater consistency in the sex-role attitudes of white women, which was first reported for the 1970 to 1973–1974 period by Mason et al. (1976), continued at least through 1975. But the level of consistency appears to have stayed the same between 1975 and 1978. We found no support, however, for the hypothesis that the increase in the consistency of sex-role attitudes for white women was part of a more general increase in the consistency of political attitudes along a liberal versus conservative dimension. Thus, whatever the merits of the argument put forward by Nie et al. (1976) and others concerning a general increase in "attitude constraint," this general phenomenon does not account for the more specific trends in the consistency of sex-role attitudes. Rather, sex-role attitudes seem to have become, for white women at least, a more distinct sphere of opinion which is increasingly structured along traditional versus nontraditional lines.

In addition, our findings suggest that the aggregate trends which have been iden-

tified for whites did not, in general, hold for blacks. Black attitudes did not change significantly between 1972 and 1978. Moreover, the associations between responses to FEWORK and FEPRES showed no clear pattern for blacks. The limited number of blacks in the GSS, however, prevented us from exploring black-white differences in detail.

We have no ready explanation for the lack of change in egalitarianism and consistency among whites after 1975, although possible reasons include greater attention in the early 1970s to the so-called "easier" feminist issues—for example, equal pay for equal work—and the mobilization of and publicity accorded to conservative groups protesting the Equal Rights Amendment and abortion since the mid-1970s. Yet despite the popular view of the mid-to-late 1970s as a time of retrenchment and conservatism, the levels of attitude egalitarianism and consistency that were attained by whites in 1975 were maintained through 1978.

REFERENCES

Barton, Allen H. and R. Wayne Parsons
1977 "Measuring belief system structure." Public Opinion Quarterly 41:159–80.

Davis, James A.
1980 "Conservative weather in a liberalizing climate: change in selected NORC General Social Survey items, 1972–1978." Social Forces 58:1129–56.

Duncan, Beverly and Otis Dudley Duncan
1978 Sex Typing and Social Roles: A Research Report. New York: Academic Press.

Ferree, Myra M.
1974 "A woman for president? Changing responses 1958–1972." Public Opinion Quarterly 38:390–99.

Haberman, Shelby J.
1978 Analysis of Qualitative Data, Volume 1: Introductory Topics. New York: Academic Press.

Mason, Karen Oppenheim, John L. Czajka, and Sara Arber
1976 "Change in U.S. women's sex-role attitudes, 1964–1974." American Sociological Review 41:573–96.

Nie, Norman H., Sidney Verba, and John R. Petrocik
1976 The Changing American Voter. Cambridge: Harvard University Press.

Smith, Tom W.
1980 A Compendium of Trends on General Social Survey Questions. Report No. 129. Chicago: National Opinion Research Center.

Spitze, Glenna and Joan Huber
1980 "Changing attitudes toward women's nonfamily roles: 1938 to 1978." Sociology of Work and Occupations 7:317–35.

Thornton, Arland and Deborah S. Freedman
1979a "Changes in the sex role attitudes of women, 1962–77: evidence from a panel study." American Sociological Review 44: 831–42.
1979b "Consistency of sex role attitudes of women, 1962–1977." Ann Arbor: Institute for Social Research, University of Michigan.

U.S. President
1979 Employment and Training Report of the President. Washington: U.S. Government Printing Office.

VARIATIONS IN SEX-RACE-AGE-SPECIFIC INCIDENCE RATES OF OFFENDING*

MICHAEL J. HINDELANG
State University of New York, Albany

Incidence rates—rates of offending in personal crimes (rape, robbery, assault, and personal larceny)—are studied using data from the National Crime Survey (NCS) for 1973–1977, conducted by the U. S. Bureau of the Census. The NCS data reveal that victims' reports of offenders' sex, race, and age are strongly related to incidence rates of offending. The highest incidence rate in personal crimes is for male, black, 18 to 20 year olds. Arrest data at the national level for robbery yield comparable results. Household crimes—burglary, household larceny, and vehicle theft—in which the victims saw and were able to report offenders' sex, race, and age constituted about 5% of all household crimes. The patterns in incidence rates of offending in these household crimes closely parallel those for personal crimes.

Until the self-report method of measuring illegal activities was used by Short and Nye (1957), most research into the factors associated with involvement in crime and deviance had relied exclusively on official arrest or offense data from the police or courts. This paper examines a third source of data, victimization survey data, for information about variations in rates of offending across demographic subgroups.

Because Short and Nye's analysis indicated that class was related to official but not self-reported delinquency (for another interpretation see Hindelang, Hirschi, and Weis, 1979), both sociological research and theory have emphasized racial and class biases in criminal justice processing. From the point of view of sociologists interested primarily in etiology, the question becomes one of measurement error introduced into official data by inappropriate labelling of persons as "delinquent" or "criminal" due solely to discrimination—that is, when there is no behavioral justification for the label—and, conversely, the failure to label an offender because of his or her privileged status.

The available research suggests that "conduct" variables (such as involvement in a crime or the seriousness of an offense) account for far more of the variance in "labelling" than do "status" variables such as race (Tittle, 1975; see Gove, 1980 for an update). For example, Wolfgang, Figlio, and Sellin (1972: Table 13.5) present a table showing the effects of their social class indicator, race, and the Sellin-Wolfgang seriousness score, on the probability that the police would release or further process boys in their cohort. The data show that (holding the remaining variables constant) on the average, lower class subjects were very slightly (4%) more likely than higher class subjects to be further processed by the police and that nonwhites were on average about 14% more likely than whites to be further processed. However, the average effect of more versus less seriousness of the instant offense yields a difference of 45% in the likelihood of further processing, with race and class controlled. This suggests negligible class bias, some racial bias, and a large effect (in the expected direction) for the seriousness of the contact offense. Many other studies produce findings of this order, indicating that, overall, although official data have labelling errors based on unwarranted discrimination, conduct differences are the principal known determinants of the label (see generally Gove, 1980).

The accuracy of official data has also been assessed by comparing the results from official data to those from self-report studies. For example, in official data the

* Direct all correspondence to Michael J. Hindelang, School of Criminal Justice, State University of New York, Albany, New York 12222.

black/white race difference in the incidence of crime and delinquency is generally large (particularly for Uniform Crime Report Index offenses), virtually irrespective of how "crime" or "delinquency" is defined (Wolfgang, et al., 1972; Webster, 1979). Unfortunately for those seeking compatibility, however, with few exceptions (e.g., Berger and Simon, 1974; Elliott and Ageton, 1980) self-report studies have found minimal differences, the mean black to white rate ratio being on the order of about 1.1:1 (e.g., Elliott and Voss, 1974; Williams and Gold, 1972; Gold, 1970; Hirschi, 1969; Epps, 1967). Although this lack of difference seems to call into question the validity of official data, self-report data, or both, there are many reasons why this conclusion may be premature. Most of the self-report studies published to date (e.g., Williams and Gold, 1972; Gold and Reimer, 1975) have had too few blacks to provide a meaningful test. In addition, omnibus or total scale scores may mask racial differences, because most self-report scales have been dominated by items that are of low seriousness and high frequency, and whites tend to report greater involvement in these events while blacks tend to report involvement in less frequent, more serious offenses (see Hindelang, Hirschi, and Weis, 1979). National victimization survey results, which focus on UCR index crimes (Webster, 1979), produce racial differences in offending that are compatible with official data (Hindelang, 1978) but are incompatible with the no-race-difference conclusion in most self-report studies. Evidence from a reverse record-check study of self-reported delinquency shows that the failure of black male offenders to report known official offenses is about three times as great as is the failure of white male offenders to report known official offenses (57% vs. 20% for serious offenses) (see Hindelang, Hirschi, and Weis, 1981). Thus, differential validity in itself may account for the failure of the self-report method to produce the racial differences expected on the basis of results from other sources (UCR arrest data and reports of victims in surveys).

One of the problems with using UCR arrest data to explore the rates of offending at the national level is the inability to construct such rates for specific population subgroups. Although the personal characteristics of arrestees are available by race and age group and by sex and age, they are not available by race and sex, nor by race, sex, and age group (this is with good reason: the data collection form used does not permit the latter two crosstabs). Furthermore, although the UCR shows total general population counts for jurisdictions reporting arrestee data, the general subpopulation details required to compute age-sex-and/or-race-specific rates of offending are not presented for the jurisdictions reporting.[1] But it is possible to circumvent these shortcomings, by studying data from the National Crime Survey.

THE PRESENT STUDY

The National Crime Survey (NCS) is a general population survey conducted by the Bureau of the Census under the specifications of the Bureau of Justice Statistics in the U. S. Department of Justice. This survey is a continuous panel survey in which nationally representative samples of households and persons are interviewed twice per year, at six month intervals throughout the year (i.e., January–June; February–July; etc.). In a six-month period, interviews are completed in about 65,000 households, in which all persons 12 years of age or older are eligible for interview (therefore, in a given six-month period more than 130,000 individuals in these 65,000 households would be interviewed). According to the NCS sponsor, "interviews were obtained in about 96% of all eligible housing units, and about 99% of the occupants of those households participating in the survey" (Law Enforcement Assistance Administration, 1976:4).

Respondents are personally interviewed and provide information about their background characteristics (such as age and education) as well as about victimizations that they (or the household) may have suffered during the previous six months.

[1] It is possible that the denominator for these rates is the same as for the United States as a whole, but this assumption may not be warranted for some subgroups.

Of particular interest here are data on the personal crimes of rape, robbery, assault, and larceny; household data will be discussed briefly below. The incidents (if any) reported to interviewers are weighted according to the inverse of the probability that they would occur in the sample. For the crimes studied here, the incident weights vary from about 1,000 to 2,000. In the analysis and presentation of data these weights have been used to provide population estimates for the United States (See Law Enforcement Assistance Administration (LEAA, 1976) for more details on procedures and the NCS instruments). The data used in this paper are national data from the 1973 to 1977 period.

Because the National Crime Survey interviewers ask victims to report on the sex, age group, and race of the offender, and because the survey generates its own age-sex-race-specific counts on the general population, it can provide age-sex-race-specific estimates of rates of offending at the national level that UCR arrest data (as currently collected) cannot.

Interviewees who report having been victims of the crimes covered by the survey are asked a set of detailed questions regarding the event, including the characteristics of the offender(s) involved. Specifically, victims are asked the age group (under 12, 12–14, 15–17, 18–20, and 21+), sex, and race of the offender(s). For the purposes of the analyses reported here, three age groups (under 18, 18–20, and 21 or older) and two racial groups (white and black) are used. Offenders of "other" races are excluded: this eliminates about 4% of the offenders in personal crimes. The data tapes provided by the Census Bureau follow its convention of classifying persons of Spanish origin as white.

Because victimization survey data are generated independently of the criminal justice system, include relatively serious offenses, and are sufficiently numerous to provide reliable estimates of rates of offending for various demographic subgroups, the NCS data seem to be worth exploring. One limitation, however, is that it is not possible to tell the extent to which a small number of offenders account for a large proportion of offenses; i.e., the survey produces incidence rather than prevalence rates. Arrest data published annually in the Uniform Crime Reports share this limitation: it is not possible to ascertain the number of distinct offenders arrested in a given period. The survey data have sufficient compensating advantages to recommend their use for studying rates of offending.

The incidence rates of offending reported in this section are based on the victims' perceptions of the offenders' age group, sex, and race. These offending-rate data are designed to parallel arrest data as closely as possible. That is, given that the survey data are incapable of providing prevalence rates based on the number of distinct offenders involved in offenses suffered by different victims, incidence rates are used. The incidence rates of offending take into account the total number of offenders in each sex-race-age subgroup subject to arrest for the offense reported to survey interviewers. For example, if one victim reports having been victimized by one white, male adult and two white, female juveniles and another victim reports having been victimized by one black, female adult and one white, male adult, the sex-race-age subtotals for these victimizations would be two white, male adults, two white, female juveniles, and one black, female adult. This subtotaling process continues across all incidents reported to survey interviewers and results in a total number of offender-weighted offenses for each sex-race-age subgroup. Rather than simply cumulating the raw numbers of offenders in each subgroup, the incident weight—the inverse of the probability that an incident will be sampled—is cumulated for each sex-race-age subgroup. This is necessary because, owing to the complex design of the survey, not every incident has the same likelihood of appearing in the sample. The subgroup totals serve as the numerators for the rates of offending reported in the following tables;[2] the denominators are

[2] We excluded incidents in which the victim did not know whether there was one or more offender because in such cases the victim was not asked the sex, race, or age of the offender(s). Incidents involving multiple offenders of mixed sexes or races were excluded for the same reason. These exclusions

estimates of the number of persons in the general population (i.e., potential offenders) in each sex-race-age subgroup.[3] Incidence rates of offending are reported per 100,000 potential offenders, and they convey the number of offender-weighted incidents occurring for every 100,000 potential offenders with specific demographic characteristics; they reflect the incidence rate of offending in personal or household victimizations.[4]

Because the oldest offender age-category in the survey data is "21 or older," it is not possible to remove from the numerator of the adult rates of offending the small proportion of crimes committed by people over, for instance, forty years of age. The result is that the adult offending rate for persons between 21 and 40 years of age is underestimated; UCR arrest data show that over 90% of arrestees for the crimes with which we are concerned are under forty. Making estimates from UCR arrest data, we increase the estimates of the incidence rates of offending for adults 21 to 40 years of age by about 60%. As will be apparent in the figures presented below, however, even if the offending rates for the adults were doubled to compensate for this phenomenon, the general patterns in the data (i.e., the adult rate of offending being the lowest) within sex-race groups would be preserved.

PERSONAL CRIMES

In the 18 to 20 and 21 or older age groups in Figure 1, black males have the highest incidence rate of offending in personal crimes (rape, robbery, assault, and larceny from the person), white males the second highest rate, black females the third highest, and white females the lowest. The pattern is similar for the 12 to 17 age group except that black females have a rate of offending slightly in excess of the rate for white males. Among females, there is a sharp, consistent decline of more than 80% in the incidence rates of offending in personal crimes as age increases. Black females in each age group are about four to five times as likely as their white counterparts to offend in face-to-face personal crimes. Among males, the black to white ratio of incidence rates of offending is also about 5:1, but the peaks in the 18 to 20 age group are striking.

The pattern of offending in Figure 1 is paralleled by the patterns in Table 1, which show the incidence rate of offending in each of the personal crimes, by sex, race, and age of offender. In this connection, violent offenses are a simple sum of rapes, aggravated assaults, and simple assaults; theft offenses are a simple sum of robberies and larcenies. Again, the peak age for violent offenses is 18 to 20 for males, where the incidence rate of offending for black males is more than three times that for white males. The incidence rates for robbery and personal larceny are relatively low for white females, black females, and white males, while the incidence rates of offending for black males are extreme outliers, eleven to twenty times those of the next highest group.

My discussion has been focused on the relative differences rather than on the absolute level of incidence rates in the data. That males, youthful (18–20) offenders,

constituted about 11% of total personal incidents. When offenders were of mixed ages, the age group of the oldest was arbitrarily used to prevent the loss of additional cases; treating mixed age-group offenders as all in the youngest age group resulted in only minor variations from the results obtained when the oldest age-group rule was used.

[3] Population bases shown in Appendix A are the estimated 1973–1977 aggregated counts in each subgroup. When these bases are divided into their corresponding 1973–1977 aggregated offender-weighted offenses, this division (times 100,000) yields the average annual rates of offending shown in the figures and tables reported here.

[4] In this paper, the general population base for the rate of juvenile offending is persons between 12 and 17 years of age. The victimization data show that fewer than one percent of the survey victimizations are committed by persons perceived by victims to have been under 12 years of age. In light of this, and since general population estimates were not made for persons under 12 because they did not fall within the scope of the victimization survey, the base of the juvenile offending rate is simply the number of 12 to 17 year olds in the general population. That is, the numerators of rates of offending for 12 to 17 year olds include a small proportion of crimes for those under 12, whereas the denominators include only persons in the general population aged 12 to 17. Because of the small number of offenders under 12, this has a trivial effect on the rates of offending for this group.

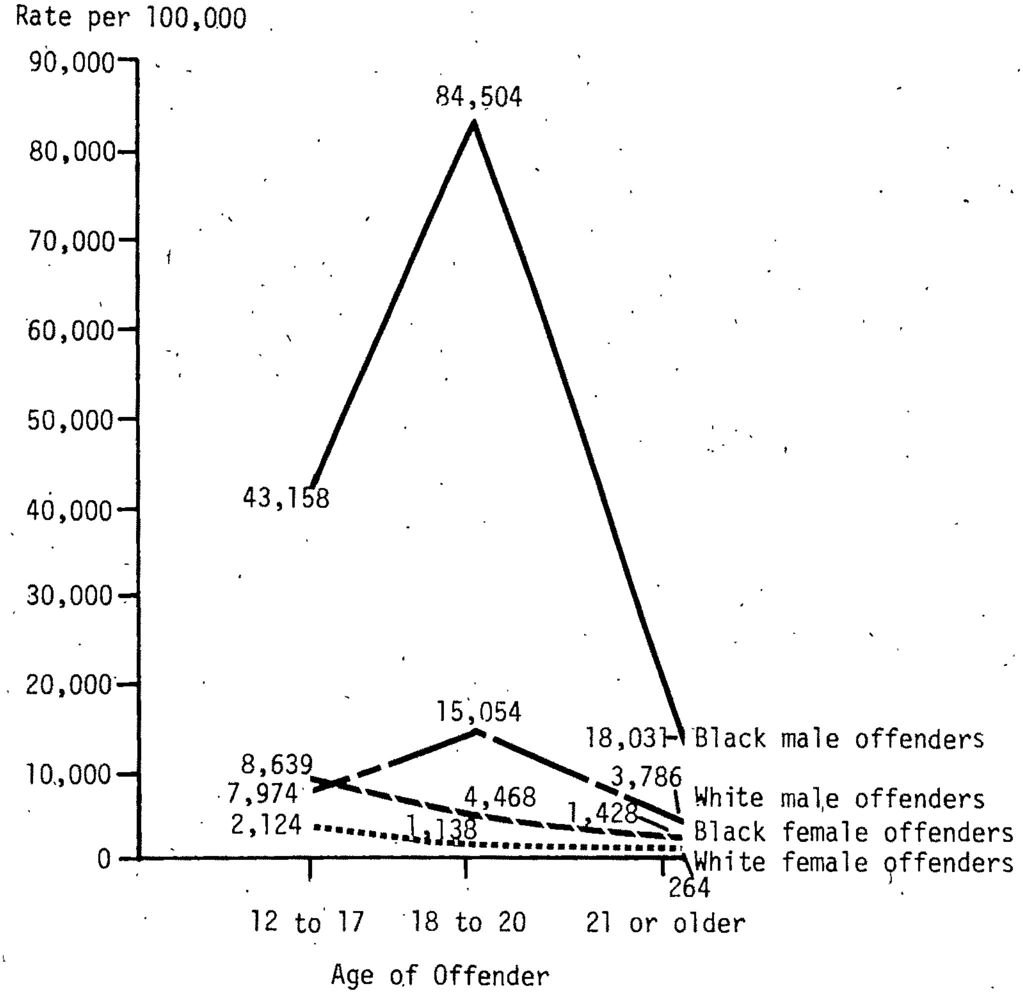


Figure 1. Estimated Annual Rates of Offending in Total Personal Crimes (per 100,000 Potential Offenders in Each Population Subgroup), NCS National Data, 1973–1977 Average

NOTE: Age includes perceived age of lone offender and perceived age of oldest member of an offending group.

Excluded are incidents (about 11% of the total) in which the victim did not know whether there was one or more than one offender and incidents involving offenders of mixed sexes or mixed races.

The numerator of the rates of offending for 12 to 17 year olds includes incidents (about 1% of the total) in which the offender was perceived by the victim to be under 12 years of age. The denominator of the rate is the number of 12 to 17 year olds in the general population. See population base estimates in Appendix A.

and blacks have the highest rates of offending could have been anticipated on the basis of prior research using official data. What is surprising, however, is the exceedingly high incidence rate of offending in personal crimes among young, black males: 85,000 offender-weighted personal offenses per year for every 100,000 persons in this subgroup. One reason why this result is so striking may be a previous lack of information: in recent years although published UCR data have presented graphs of the age distributions of the general population, no data taking race, sex, and age group into account simultaneously have been presented (Kelley, 1977:171). Also, any rates derived from UCR data are based on arrests, but an arrest is unlikely if the offense is not reported to the police (about half of personal crimes reported to survey interviewers are not reported to the police). Even if reported, the crime must be "solved" before an arrest can be made. Each of these conditions, while perhaps not affecting the relative rates of offense

Table 1. Estimated Annual Rates of Offending in Personal Crimes (per 100,000 Potential Offenders in Each Population Subgroup), NCS National Data, 1973–1977 Aggregate

Race, Sex and Age of Offender	Rape	Robbery	Aggravated Assault and Simple Assault	Personal Larceny	Violent	Theft	Total
White males:							
12 to 17[a]	77	1,203	6,436	259	6,513	1,462	7,974
18 to 20	291	2,245	12,144	374	12,435	2,619	15,054
21 or older	152	463	2,100	72	3,252	535	3,786
Black males:							
12 to 17[a]	403	16,663	21,619	4,473	22,022	21,136	43,158
18 to 20	1,624	35,030	39,990	7,860	41,614	42,890	84,504
21 or older	735	7,000	9,108	1,188	9,843	8,188	18,031
White females:							
12 to 17[a]	5	212	1,854	52	1,859	264	2,124
18 to 20	0	71	1,035	33	1,035	104	1,138
21 or older	0	33	219	11	219	44	264
Black females:							
12 to 17[a]	92	1,307	6,736	504	6,828	1,811	8,639
18 to 20	39	703	3,074	651	3,113	1,354	4,468
21 or older	7	164	1,120	138	1,127	302	1,428

NOTE: Age includes perceived age of lone and perceived age of oldest multiple offender.

Excluded are incidents (about 11 percent of the total) in which the victim did not know whether there was one or more than one offender, and incidents involving offenders of "mixed" sexes or "mixed" races.

[a] The numerator of the rates of offending for 12 to 17 year olds includes incidents (about 1 percent of the total) in which the offender was perceived by the victim to be under 12 years of age. The denominator of the rate is the number of 12 to 17 year olds in the general population. See population base estimates in Appendix A.

significantly, reduces the absolute rates substantially.

It must be recalled that the rates of offending presented are not prevalence rates; there are an estimated 84,504 incidents committed by offenders in personal crimes—many of whom are duplicative—per 100,000 persons in this sex-race-age subgroup. Part of the problem may be that this high-rate group is the 18 to 20 year old age group, for whom victims' perceptions are probably more prone to error, due to the narrowness of the age category. This would suggest looking to the extreme age groups for greater stability in comparisons. But even among juveniles (12–17), an age category encompassing six rather than three years, and hence presumably less subject to error in victim perceptions, the incidence rate of offending for black males is 43,158 per 100,000, a very high annual rate, and one which is five times greater than that of their white male counterparts (7,974).

How can we determine whether these rates are on the order of magnitude that we would expect on the basis of other data? UCR national arrest data seem appropriate as a comparison source, especially the data on robbery.[5] One important question is to what extent the percent distribution of UCR robbery arrestees for a given year approximates that for NCS robbery offenders across age group and race. Because, with the exception of offenders' age, UCR data are more limited than NCS data in terms of demographic subgroups for which national data are available, the UCR data-presentation format will limit the comparisons that can be made. The data in Table 2 show, for age group and race, the percent distribution of all UCR 1976 robbery arrestees falling into each race and age group cell; comparable data are shown for NCS offender-weighted incidents. The two data sources are in close agreement regarding these percentages. For example, UCR arrest data and NCS offender data show that about one-fifth of the robberies were accounted for by black juveniles, and yet

[5] The crime of rape is too rare statistically to use it. Historically, the NCS has had some difficulty measuring assault (LEAA, 1972) and the UCR larceny category includes arrests for many larcenies not in the scope of the NCS, particularly shoplifting. The NCS robbery data used in the comparison to the UCR exclude commercial robberies.

Table 2. Percent Distributions for Offender-Weighted Incidents in NCS Data and Arrests in UCR Data, for Personal Robbery, 1976 National Data

Age of Offender	Race of Offender: White	Black	Total
Under 18			
UCR	11[a]	20	31
	(9,728)	(17,929)	(27,657)
NCS	11[b]	19	30
	(176,901)	(297,972)	(474,873)
18 or older			
UCR	29	40	69
	(25,500)	(35,752)	(61,252)
NCS	26	44	70
	(402,506)	(688,205)	(1,090,711)

SOURCE: UCR data are from Kelley, 1977:186, 187.

[a] Percent based on total number of UCR robbery arrests in 1976 as shown in Kelley, 1977:186, 187. "Other" races are excluded.

[b] Percent based on total estimated number of NCS offender-weighted robbery incidents. Excluded are incidents (about 8% of the total) in which the victim did not know whether there was one or more than one offender and incidents involving offenders of mixed races. Other races are also excluded.

black juveniles constitute only 2% of the general population; this demographic subgroup, therefore, is represented among robbers at about 10 times their proportionate representation in the general population. When age group and sex are examined in a similar way, both the UCR and NCS data (not shown in tabular form) indicate that in 1976 about three out of ten robbery arrestees (UCR) or offenders (NCS) were males under 18 years of age. This is a substantial (4:1) overrepresentation of the proportion of 12 to 17 year old males in the general population.

The UCR and NCS data can be said to produce strikingly similar distributions for these variables, considered on the bivariate basis shown. Unfortunately, the UCR does not present the requisite trivariate age, sex, and race arrest data required to check further the NCS offending-rate results. However, by making an assumption we can estimate the trivariate distribution among UCR robbery arrestees by age group, sex, and race.

From UCR data on arrestee's age group, sex, and race (e.g., Kelley, 1977: Tables 32, 34, and 35) we can obtain at least a crude approximation of arrestee's trivariate age-group-sex-race distribution by assuming no correlation between demographic variables whose joint distributions are not published in the UCR (e.g., sex by race). Undoubtedly, this introduces some error, but because these estimates are for order-of-magnitude purposes, we need not be unduly concerned at this point. By way of illustration, UCR arrest data show that 93% of robbery arrestees are male, that 23% are 18 to 20 years of age, and that 60% are black. In Table 3, therefore, the percentage of all UCR arrestees estimated to be male, 18 to 20 years of age, and black is 93% × 23% × 60% = 12.8%. Table 3 shows (Column A) the estimated UCR arrest percent distribution across all age-sex-race cells; also shown (Column B) is the percent distribution of NCS offender-weighted robbers. As was true at the bivariate level, the UCR and NCS data at the trivariate level are in general agreement. Both UCR and NCS percent distribution data indicate that white females and adults tend to be underrepresented among robbery offenders in comparison to the representation of these demographic subgroups in the general population.

Columns D and E index the overrepresentation of demographic subgroups for the UCR and NCS data, respectively. As was foreshadowed by the data in Columns A and B, the sources are again similar. Here, however, the ratio of the percent distribution in the offending populations to the percent distribution in the general population is a clear index of the extent of the overrepresentation of particular subgroups. Within Columns D and E, the index figures can be viewed in relation to each other. In both columns, for example, black males 18 to 20 years of age are overrepresented among robbery offenders at the rate of 29 or more to 1. At the other extreme, white females of all ages tend to be substantially underrepresented among offenders by either the UCR or NCS criterion.

If we assume that the victimization survey data provide our best estimate of the total number of offender-weighted personal robberies—primarily because the NCS data show that only half of these robberies become offenses known to the

Table 3. Comparison of NCS and UCR Offense Data for Personal Robbery, 1976 National Data

Race, Sex, and Age of Offender	A Estimated percent distribution of UCR robbery arrestees[a]	B Estimated percent distribution of NCS offender-weighted robbers[b]	C Estimated percent distribution of the general population[c]	D Ratio of UCR percent of arrestee population to percent of general population	E Ratio of NCS percent of robber population to percent of general population	F Estimated rate of robbery offending (per 100,000) based on UCR percents (Column A) times estimated NCS total of offender-weighted robbers divided by general population count for each subgroup	G Estimated rate of robbery offending (per 100,000) based on NCS percents (Column B) times estimated NCS total of offender-weighted robberies divided by general population count for each subgroup	H Ratio of estimated UCR rates (See Column F) to estimated NCS rates (See Column G)
White males:								
12 to 17	12.33	9.68	6.25	1.96	1.55	1,706	1,340	1.27
18 to 20	8.42	6.85	3.00	2.81	2.28	2,432	1,980	1.23
21 or older	16.00	17.81	33.57	.48	.53	413	459	.90
Black males:								
12 to 17	18.79	18.19	1.04	18.07	17.49	15,615	15,116	1.03
18 to 20	12.84	15.94	.44	29.18	36.23	25,263	31,358	.81
21 or older	24.39	26.38	3.57	6.83	7.39	5,919	6,400	.92
White females:								
12 to 17	.96	1.07	6.01	.16	.18	138	154	.90
18 to 20	.66	.34	3.10	.21	.11	184	95	1.94
21 or older	1.25	1.10	36.98	.03	.30	29	26	1.12
Black females:								
12 to 17	1.46	.74	1.05	1.39	.70	1,203	613	1.97
18 to 20	1.00	.74	.51	1.96	1.45	1,704	1,254	1.35
21 or older	1.90	1.16	4.48	.42	.26	367	225	1.63

[a] UCR data are from Kelley, 1977; Tables 32, 34, and 35. Trivariate distribution is estimated based on the data in Kelley (1977: Tables 32, 34 and 35) under the assumption that the demographic variables are uncorrelated with each other. Note "other" races are excluded.

[b] These data are taken from the NCS survey for 1976. Excluded are incidents (about 11 percent of the total) in which the victim did not know whether there was one or more than one offender and incidents involving offenders of "mixed" sexes or "mixed" races. Note "other" races are excluded.

[c] Estimates based on NCS data for 1976. Persons under 12 are ineligible to be interviewed and are excluded from the survey.

police and a relatively small percentage of those reported result in arrest (Webster, 1979)—we can use this base to convert the percentage figures shown in Columns A and B to offender-weighted rates of personal robbery victimizations for the UCR and NCS percentages, Columns F and G. To the extent that the respective percentages in Columns A and B are similar, so too are the corresponding rates of personal robbery offending. The level of rates in Columns F and G, however, are a function not only of the percentages shown in A and B, but also of the total number of offender-weighted robberies against which these percentages are applied. Using the NCS total number of offender-weighted personal robberies as the best available estimate of persons subject to arrest for committing personal robberies will not bias the resultant incidence rates of offending; it is, in effect, a linear transformation of the UCR and NCS percentages.

As expected in light of the data in Columns A through E, the incidence-rate data in Columns F and G are compatible with each other. The rates for black males in all age groups—particularly the two youngest groups—stand out from those of all other groups. This is true whether the UCR or the NCS criterion is used. For example, according to UCR-based estimates, black males 18 to 20 and juvenile black males have incidence rates of robbery offending per 100,000 of 25,263 and 15,615; the comparable rates in the NCS are larger for the youthful offenders and slightly smaller for juvenile offenders (31,362 per 100,000 and 15,116 per 100,000). The similarity in the rankings of rates is indicated by a Spearman rank-order correlation coefficient of .95.

In summary, the extreme incidence rates of robbery for black male youth are not caused primarily by the NCS count being greater than the UCR arrest count. Plotting Column D's values for white and black males will produce a graph very much like that shown in Figure 1.

HOUSEHOLD CRIMES

In addition to the personal crimes discussed above, the NCS collects data on crimes construed to affect the household—burglary, household larceny, and vehicle theft. These data are collected from the "household respondent," someone who is knowledgeable about the affairs of the household. Because of the nature of these crimes—typically the victim and the offender(s) do not come face to face—for most of them offender's characteristics are unknown. For a small minority of household crimes, however, the victim and offender do come into contact. In the 1973–1977 period, victims were able to report on offenders' characteristics in 6% of the burglaries and vehicle thefts and in 4% of the household larcenies. Because of the large number of household crimes occurring in this period, even these small proportions yield large numbers of events to analyze: 6,000 unweighted (sample) household crimes reported, 1,600 burglaries, 4,000 larcenies, and 300 vehicle thefts—or an estimated (weighted) 2 million burglaries, 5 million larcenies, and .4 million vehicle thefts. There is no a priori reason to believe that there are offender-linked characteristics associated with those household crimes that do (about 5%) or do not (about 95%) result in the victims' being able to report offenders' characteristics to the interviewer.[6] Because the clearance rate for these crimes in UCR data is so low—less than 20% (Kelley, 1977:161)—when the rate at which NCS incidents are reported to the police is taken into account [about 70% for vehicle theft, 50% for burglary, and 25% for household larceny, according to NCS figures (LEAA, 1976)], UCR arrest data also exclude the vast majority of offenders in these crimes. Therefore, NCS data may be superior to UCR arrest data for analysis of the characteristics of household-crime offenders because the NCS data are closer to the crime (Sellin, 1951) and subject to fewer filtering mechanisms than the UCR arrest data.

Figure 2 depicts very striking sex, race, and age group differences in incidence rates of offending for those household of-

[6] Household crimes in which the offender(s) were seen, as expected, were more likely to be reported to the police (e.g., burglary: 61% versus 50%) and less likely to result in something being stolen (all household crimes: 52% versus 79%).

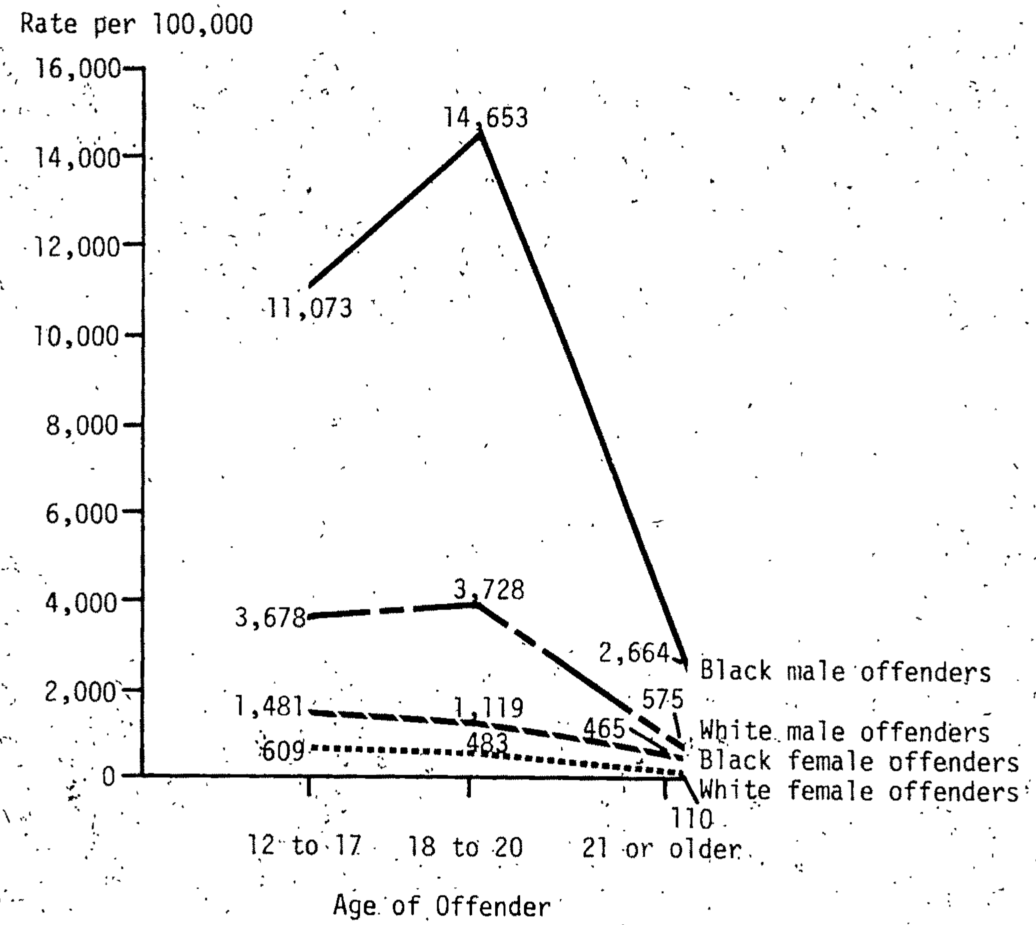


Figure 2. Estimated Annual Rates of Offending in Total Household Crimes in Which the Offender Was Seen (per 100,000 Potential Offenders in Each Population Subgroup), NCS National Data, 1973–1977 Average

NOTE: About 95% of household crimes, in which the offender was not seen, are excluded.
Age includes perceived age of lone offender and perceived age of oldest member of an offending group.
Excluded are incidents in which the victim did not know whether there was one or more than one offender and which involved offenders of mixed sexes or races.
The numerator of the rates of offending for 12 to 17 year olds includes incidents (about 1% of the total) in which the offender was perceived by the victim to be under 12 years of age. The denominator of the rate is the number of 12 to 17 year olds in the general population. See population base estimates in Appendix A.

fenses in which victims reported offenders' demographic characteristics; these patterns mirror those shown for incidence rates of offending in personal crimes. When the individual household crimes are examined (data not presented) they show, for example, that the incidence rate of burglary offending for the black, male, 18–20 year old group (4,446) is more than 5 times the rate of the next closest white sex-age-group (white males, 12 to 17: 871 per 100,000). Again, among black males, the incidence rate of burglary offending shows a sharp rate increase from the juveniles (1,985) to the youthful offender group (4,446), which then falls sharply to the (still relatively quite high) incidence rate of more than (800) for black male adults. In the three age groups, the smallest black-white difference is 2:1 among juveniles and the largest is among adults (about 5:1). The other household crimes show similar patterns to that for burglary, and hence for the personal crimes examined above.

DATA LIMITATIONS

The NCS survey only includes data on "street" crime of the UCR index variety.

The NCS is also known to undercount all offenses, particularly assault (LEAA, 1972) and to be subject to certain time-in-panel biases which reduce reported victimizations (Woltman and Bushery, 1977). Most important to this study is that the validity of results relies entirely on victims' perceptions of the characteristics of the offenders, and to the extent that these are inaccurate, the analysis must necessarily fall short. However, victims could (and sometimes did) answer that they "did not know," and these incidents were excluded from the analysis, which should have reduced biases in the data set.

While we would not expect the accuracy of victims' reports of the sex of the offender to be problematic, reports of the offender's race might well be. For example, persons of Spanish origin might be mistaken by some victims as black, but this would present difficulties in only some sections of the country. Age was expected to be the variable most difficult for victims to estimate, and so the NCS data were separated into three broad groups for the purpose of analysis and presentation; when age is used as a dichotomy (12–20 versus 21 or more) the results are not substantially affected, and when age is dichotomized the other way (12–18 versus 18 or more) the age effect diminishes but it is still pronounced.

In order to test the accuracy of victims' reports on offenders' characteristics, data available from a special study of rape in New York City were analyzed.[7] The study involved about 13,000 offenses for which data on suspects were collected; for most suspects (about 8,000), no arrest was made. For the 5,000 offenses in which an arrest was made, descriptions of the demographic characteristics of suspects reported to the police at the time of the offense were crosstabulated with the arrestees' characteristics. The age categories available from those data are under 14, 14 to 19, 20 to 24, 25 to 29, 30 to 34, 35 to 39, 40 to 45, 45 or more, and "don't know." Of the suspects reported to be under 14, 97% were found to be under 14 when an arrest was made. The following age groups estimated at the time of the offense yielded precisely the same age group at arrest: 14 to 19, 96%; 20 to 24, 89%; 25 to 29, 90%; 30 to 34, 90%; 35 to 39, 89%; 40 to 45, 91%; 45 or more, 96%. The offenders' age group reported on the offense form and the arrestees' age group yielded a Somers' d of .95. It is important to note that the agreement does not decline appreciably when only victims and offenders who were strangers are analyzed.

The race/ethnicity categories available in the New York data include persons of Spanish origin and hence are finer than the categories available for analysis in the NCS; they provide a more stringent test of victims' ability to report accurately on offenders' race.[8] The agreement between reports on the offender's race at the time the offense was reported and the race of the arrestee recorded on the arrest form was very high (lambda = .95). Among offenders reported to be white at the time of the offense, 96% were recorded as white on the arrest form. For blacks, the comparable figure is 99%, and for Hispanics, 98%. Of particular interest is the finding that only 1.6% of the Hispanic arrestees were erroneously reported to be black at the time of the offense (Hindelang and McDermott, 1981). These data suggest that, at least for this crime, victims are very accurate in their reports of offenders' age and race/ethnicity. But there are features to rape that may limit the generality of these findings, particularly the physical contact between the victim and the offender, which is sometimes of an extended nature in completed rapes, and the fact that the sex of the victim is virtually always female (99%). If we assume that attempted rapes are of a shorter duration than completed rapes, the former might be expected to produce less accurate offender data; however, the Somers' d value for offenders' age in attempted rapes as reported at the time of the offense and at the time of the arrest is .96, while the

[7] The data were provided by Sergeant Dennis Butler of the New York City Police Department. We are grateful for his assistance.

[8] As noted for race in the NCS, the Census Bureau's classification system counts Hispanic and Anglo offenders as white; therefore, even if their numbers had been sufficiently large, they could not have been treated separately in the analysis of offending rates.

corresponding measure for race yields a lambda of .96, coefficients as high as in completed rapes. Despite the equality of the coefficients for attempted and completed rapes, these levels of accuracy should certainly be considered tentative until replicated with an independent data set.

DISCUSSION

The findings presented here in the form of incidence rates of offending parallel the findings in most sources of arrest data, despite sociologists' reluctance in recent decades to accept police and court data as evidence of involvement in crime. In addition, presenting the data in the form of the rates highlights most dramatically how closely both the crimes of adults and the crimes of juveniles are associated with the offenders' demographic characteristics. This is true not only for robbery, the "test" offense used in Table 3, but for all of the other personal and household offenses studied as well.

Very few data sets in the study of crime include information on adult offenders as well as juveniles, and certainly not of the average level of seriousness of the offenses reported to the NCS survey.[9] This lack of data, in combination with the fact that adolescent in-school populations are easier to study than general adult populations, has had the consequence that in sociology, etiological theories of juvenile delinquency are more common than etiological theories of adult crime. If we accept the NCS results at face value, they have important implications for sociological theory. The age, race, and sex variables account for an impressive percentage of the variability in incidence rates of offending. The lowest rate group, white, female adults, has an incidence rate of offending in personal crimes of less than 270, while the black, male, 18 to 20 year old group has an incidence rate of offending that is more than 300 times this rate. (If the smaller base of adults in the offending age range is used, this ratio drops to about 150.) The ratio between the incidence rates of white, male adults and 18–20 year old, black males is 22:1. The sorts of variables used by contemporary sociological theorists to explain involvement in criminal and delinquent behavior—social class, differential opportunity, stakes in conformity, maturational reform, educational aspirations, pluralistic ignorance, attachments, etc.—are rarely found in research to produce ratios in rates of offending between contrast groups (in which each contains a reasonable proportion of the sample) that exceed two or three to one (Williams and Gold, 1972; Gold and Reimer, 1975; Elliott and Ageton, 1980).

It is likely, of course, that the demographic variables studied here are, themselves, indicators of variables like these that are more likely to be found in sociological theories of etiology. For example, late adolescence (18–20), in the market place of employment, may be the time when the effects of differential opportunities are greatest and attachments to the school and family are the least; hence it is the peak age of offending among males. In this age group, blacks particularly may have lower stakes in conformity due to their historically high rate of unemployment and an average educational attainment level that is lower than that of whites; blacks correctly perceive that they have lower stakes in conforming to legal norms. The large sex difference may be due to the greater attachment of females than males to each other, to the school, to their families of origin, and later to their offspring. The much lower rate of offending for adults than for the two youngest age groups may be indicative of maturational reform. Unfortunately, the nature and pattern of demographic correlates of crime are often compatible with competing theories.

[9] Despite the fact that as a set the NCS data contain many serious victimizations in terms of harm to the victim—certainly events that have a median harm value in excess of national self-report surveys (e.g., Elliott and Ageton, 1980)—the NCS data also contain many reports of victimizations that would have to be characterized as minor. For example, of NCS victimizations categorized using the Sellin and Wolfgang (1964) system, Gottfredson and Hindelang (1979) found one-quarter resulting in no harm or minor harm (e.g., scrapes) to the victim. Thus, the NCS data cover a very wide range of seriousness, and yet as a set simply include more serious offenses, largely because of the huge number of personal interviews conducted each month.

The question of what specific mechanisms link particular demographic variables to offensive behavior must be addressed by research beyond the scope of the NCS data. Regardless of what those speculations are, these data strongly support the importance of sex, race, and age in accounting for differences in rates of offending. The strength of these correlations with criminal behavior was anticipated from arrest data (e.g., Kelley, 1977: Tables 32, 34, 35). However, the general agreement between the UCR and NCS on offenders' sex, race, and age characteristics increases the probability that both are acceptably valid for the purposes of this paper, particularly because the NCS data are collected independently of the criminal justice system and hence are subject to very different distortions than arrest data.[10]

If sociological theorists of crime and delinquency were to use the "clues" provided by known correlates of criminal behavior—in this instance sex, race, and age group—as a basis for generating and modifying theory, theory and research might be able to advance more steadily. Sociological theorists must demonstrate their ability to accommodate these associations of demographic variables to incidence rates of offending in "street" crime before their theories are taken seriously. Theories that cannot should be discounted until they can.

[10] Ecological-level comparisons of the UCR/NCS correlation for rates of crime found that the two sources are in general agreement for theft crimes, but that for some other crimes, particularly assault, the intercity correspondence across 26 cities is poor (Nelson, 1979). Although these assault findings suggest that caution is warranted, they do not necessarily say anything about the individual-level agreement between the two sources in, for example, offender characteristics. It is possible, in other words, for both sources to provide reasonably accurate data on offender characteristics and yet, for a variety of reasons, have uncorrelated intercity rates of offenses. In short, factors that vary from city to city and serve to over- or underestimate rates of crime as measured in either source do not necessarily distort characteristics of offenders as measured by the two methods. Likewise, on the national level, comparisons of UCR and NCS trends (Eck and Riccio, 1979) say little in themselves about the validity of correlates of offending on the individual level.

Appendix A

Estimated Population Bases by Age, Race, and Sex, NCS National Data, 1973–1977 Aggregate

	Age		
Sex and Race	12 to 17	18 to 20	21 or older
Male	62,066,784	28,383,692	309,537,275
White	53,287,440	24,849,625	279,948,590
Black	8,779,344	3,534,067	29,588,685
Female	60,127,648	30,120,387	345,710,930
White	51,328,096	25,970,146	308,668,220
Black	8,799,552	4,150,241	37,042,710

NOTE: Population bases shown are approximately five times as large as those for any single year.

REFERENCES

Berger, Alan S. and William Simon
1974 "Black families and the Moynihan report: a research evaluation." Social Problems 22:146–61.

Eck, J. and Lucius Riccio
1979 "Relationship between reported crime rates and victimization survey results." Journal of Criminal Justice 7:293–308.

Elliott, Delbert and Suzanne Ageton
1980 "Reconciling differences in estimates of delinquency." American Sociological Review 45:95–110.

Elliott, Delbert S. and Harwin L. Voss
1974 Delinquency and Dropout. Lexington, Mass.: D. C. Heath.

Epps, E. G.
1967 "Socioeconomic status, race, level of aspiration, and juvenile delinquency: a limited empirical test of Merton's conception of deviation." Phylon 28:16–27.

Gold, Martin
1970 Delinquent Behavior in an American City. Belmont, Ca.: Brooks/Cole.

Gold, Martin and David J. Reimer
1975 "Changing patterns of delinquent behavior among Americans 13 through 16 years old: 1967–1972." Crime and Delinquency Literature 7:483–517.

Gordon, Robert A.
1976 "Prevalence: the rare datum in delinquency measurement and its implications for the theory of delinquency." Pp. 201–89 in Malcolm Klein (ed.), The Juvenile Justice System. Beverly Hills, Ca.: Sage Publications.

Gottfredson, Michael R. and Michael J. Hindelang
1979 "A study of the behavior of law." American Sociological Review 44:1–28.

Gove, Walter
1980 "Postscript to labelling and crime." Pp. 264–9 in Walter Gove (ed.), The Labelling of Deviance, 2nd ed. Beverly Hills: Sage Publications.

Hindelang, Michael J.
1978 "Race and involvement in common law personal crimes." American Sociological Review 43:93–109.

Hindelang, Michael J., Travis Hirschi, and Joseph G. Weis
1979 "Correlates of delinquency: the illusion of discrepancy between self-report and official measures." American Sociological Review 44:955–1014.
1981 Measuring Delinquency. Beverly Hills: Sage Publications.
Hindelang, Michael J. and M. Joan McDermott
1981 Juvenile Criminal Behavior: An Analysis of Rates and Victim Characteristics. Washington, D. C.: Government Printing Office.
Hirschi, Travis
1969 Causes of Delinquency. Berkeley: University of California Press.
Kelley, Clarence
1977 Crime in the United States. Washington, D.C.: U. S. Government Printing Office.
Law Enforcement Assistance Administration
1972 San Jose Methods Test of Crime Victims. Statistics Technical Report No. 1. Washington, D. C.: U. S. Government Printing Office.
1976 Criminal Victimization in the United States. Washington, D. C.: U. S. Government Printing Office.
Nelson, James F.
1979 "Implications for the ecological study of crime: a research note." In W. H. Parsonage (ed.) Victimology. Beverly Hills: Sage Publications.
Schmid, C. F.
1960a "Urban crime areas: part I." American Sociological Review 25:527–42.
1960b "Urban crime areas: part II." American Sociological Review 25:655–78.
Sellin, Thorsten
1951 "The significance of records of crime." Law Quarterly Review 67:489–504.
Sellin, Thorsten and Marvin Wolfgang
1964 The Measurement of Delinquency. New York: Wiley.
Shaw, Clifford R. and Henry D. McKay
1942 Juvenile Delinquency and Urban Areas. Chicago: University of Chicago Press.
Short, James F. and F. Ivan Nye
1957 "Reported behavior as a criterion of deviant behavior." Social Problems 5:207–13.
Tittle, Charles R.
1975 "Labelling and crime: an empirical evaluation." Pp. 241–70 in Walter Gove (ed.), The Labelling of Deviance: Evaluating a Perspective, 2nd ed. Beverly Hills: Sage.
Webster, William H.
1979 Crime in the United States, 1978. Washington, D. C.: Government Printing Office.
Williams, Jay R. and Martin Gold
1972 "From delinquent behavior to official delinquency." Social Problems 20:209–29.
Wolfgang, Marvin, Robert Figlio, and Thorsten Sellin
1972 Delinquency in a Birth Cohort. Chicago: University of Chicago Press.
Woltman, Henry and John Bushery
1977 "Update of the NCS panel bias study." Mimeographed. U. S. Bureau of the Census, Washington, D. C.

A DYNAMIC STUDY OF CHANGE IN A REGIONAL CORPORATE NETWORK*

JOSEPH GALASKIEWICZ
University of Minnesota, Minneapolis

STANLEY WASSERMAN
University of Minnesota, St. Paul

We adapt a class of new stochastic models for social networks to the study of social change in corporate interlock networks. Data on a regional (Minnesota) network are used to verify several descriptive hypotheses drawn from the existing literature concerning interlocking directorates. We conclude that corporations are more likely to make reciprocal board linkages asymmetric than they are to reciprocate asymmetric ties, popular firms are just as likely as less popular firms to be recruited to new boards, popular firms are more likely than less popular firms to leave boards, and nonfinancial business organizations are more likely to form new board ties with commercial banks and insurance companies than with other nonfinancial firms.

In this research note we demonstrate how a class of stochastic models developed by Holland and Leinhardt (1977a, 1977b) and Wasserman (1979, 1980) can be used to test four simple descriptive propositions that are drawn from the managerialist (Allen, 1974; Pfeffer, 1972; Herman, 1973; Mace, 1971; and Burt, 1979) and finance capitalist (Perlo, 1957; Fitch and Oppenheimer, 1970; and Mintz and Schwartz, 1977) literatures on corporate interlocks. One problem which we and others (e.g., Fennema and Schijf, 1979) have encountered is that propositions from these literatures either are only relevant to cross-sectional data or, if they include a temporal dimension, have ignored network properties of corporate interlock systems. From the large number of propositions that could be derived from these two perspectives, we consider only those where it is reasonable to assume that changes in network properties can be described as stochastic processes, specifically Markov chains. With longitudinal data on networks of interlocking corporate interlocks in the state of Minnesota from 1969 through 1978, we show how stochastic models can be used to explore substantive interorganizational network change.[1]

The Evolution of Corporate Interlock Networks: Some Hypotheses

Managerialists argue that directors are important channels through which information flows between firms. For example,

* Direct all correspondence to Joseph Galaskiewicz, Department of Sociology, University of Minnesota, Minneapolis, Minnesota 55455 or Stanley Wasserman, School of Statistics, University of Minnesota, St. Paul, Minnesota 55108.

Authorship is in alphabetical order. This research was supported in part by National Science Foundation Grants SOC78-26075 and SES80-08573. Additional support was provided by the Social Science Research Council and a Grant-in-Aid of Research from the University of Minnesota Graduate School. Data collection supported by funds from the Department of Sociology, University of Minnesota. We would like to thank Barbara Rauschenbach and Wayne Kobbervig for research assistance and Phipps Arabie, Ronald Burt, John Delany, and the referees for useful comments on earlier drafts. A version of this paper was presented at the American Sociological Association 1979 Annual Meeting, Session on Social Networks, Boston, Massachusetts, August 1979.

[1] Our conceptualization of a "corporate interlock" is slightly different from that currently found in much of the literature. As defined by Allen (1974:393), ". . . the term 'interlocking directorate' typically refers to any situation in which two or more corporations share one or more directors in common." We seek to restore asymmetry to board linkages (see Fennema and Schijf, 1979; and Warner and Unwalla, 1967) and thus define an interlock between firms as any situation in which an officer of one firm sits on the board of another firm. Although there is merit in specifying indirect as well as direct ties (see Berkowitz, et al., 1979), only direct ties through corporate officers are considered.

in his study of corporate boards, Mace (1971:13) described how a host firm often solicits information and advice from outside board members. Mace (1971) also described how board members, in return, are often privy to inside information. Extensive information exchange could give rise to conflicts of interest, violations of SEC regulations, and opportunism in the marketplace; nevertheless, information exchanges between board members and their host firms appear to be quite common.

Given that information is freely passed back and forth between firms through a single director, we hypothesize that board linkages between firms would not be reciprocal. If firm j is represented on firm i's board, then there is really no need for firm i to be represented on firm j's board as well. This hypothesis is simple enough. Because it is costly to use up a board slot on a firm to which one is already linked, firms should refrain from establishing reciprocal board linkages and terminate reciprocal linkages where they do exist.

> Hypothesis 1: The probability of reciprocated board linkages becoming nonreciprocated is greater than the probability of nonreciprocated board linkages becoming reciprocated.

Because board members are important information conduits, we also expect that the number of boards on which a firm is represented will affect its likelihood of recruitment to different corporate boards in the future. If a firm is on several boards, it is in the center of an important resource network and has access to valuable information. Consequently, its officers should be very attractive potential board members. To take advantage of this access to information, firms are likely to selectively recruit those board members who have the most to contribute. Thus, firms that are popular tend to become even more popular.

> Hypothesis 2: If firm A is represented on more boards than firm B, at time t, then the probability of one or more new representations at t+1 will be greater for A than for B.

We expect, though, that there is an upper limit to the number of boards on which a firm can be represented. Even though a firm's officers may be in high demand, a firm has a limited number of officers whose time must be allocated to many activities. One solution is simply not to go on any more boards; another solution is to attenuate present linkages and to take advantage of new opportunities. We believe that firms choose the latter. To improve the information to which they have access, firms might find it difficult to resist an invitation to expand their contacts.

> Hypothesis 3: If firm A is represented on more boards than firm B at time t, then the probability of leaving one or more old representations at time t+1 will be greater for A than for B.

Finance capitalists argue that, as capitalism matures, banks and insurance companies extend their control over corporations. Indeed, financial institutions have consistently been found at the centers of corporate interlock networks (see Dooley, 1969; Mariolis, 1975; Allen, 1974; Sonquist and Koenig, 1975). As Mintz and Schwartz (1977) argue, financial organizations are the hubs of industrial organizations. As controllers of capital, these organizations are vital to the interests of large corporations. They are potential sources of funds for internal expansion, product diversification, and acquisitions. Consequently, we expect that vacant board slots among nonfinancial firms are more likely to be filled by either a commercial bank or insurance company than another nonfinancial firm. In addition, as officers of nonfinancial companies go onto new boards, they should be more likely to go onto the boards of banks and insurance companies.

> Hypothesis 4: As industrial, retail, utility, and transportation firms form new board linkages, the probability that these new linkages are with other industrial, retail, utility, and transportation firms is smaller than the probability that these new linkages are with commercial banks and insurance companies.

However, this may be an oversimplification, since an economic system does not evolve without fluctuation or variation. For example, in a period of acquisitions (see O'Connor, 1968) or just before and

after a slump in the business cycle (see Fitch and Oppenheimer, 1970), ties to financial institutions become more centralized (e.g., in New York banks and insurance companies) and corporations turn to extralocal sources for loans. Under these circumstances we suspect that local corporations would withdraw from board interlocks with local financial institutions (see Mintz and Schwartz, 1977:22). Needless to say, Hypothesis 4 ignores the business cycle during a period of great instability. Our only defense is that the economy of the Upper Midwest (which is based on high technology and foodstuffs) has prospered in the period studied and traditionally has not suffered as much during recessions.

Research Site and Data

The network under study consists of twenty-six of the largest public business corporations in Minnesota from 1969 to 1978. We examined *Fortune* magazine's listing of the 1,000 largest industrial firms, 50 largest retail firms, 50 largest transportation firms, 50 largest utilities, 50 largest commercial banks, and 50 largest life insurance companies for each year and identified all business firms located in Minnesota. We then chose for our study only the firms that made the *Fortune* lists each of the ten years.[2]

From annual reports we identified board members and their primary business affiliation. We constructed for each year a binary 26 × 26 adjacency matrix. A "1" was placed in a cell (i,j) if an officer of organization j was on the board of corporation i, and a "0" otherwise. We ignored retired officers. Diagonal elements of the matrices were also ignored.

Several very large corporations were omitted from our study by design. Those firms that did not appear on the *Fortune* lists for ten consecutive years were dropped.[3] Thus, firms that had recently experienced rapid growth or decline, new firms, and large firms that had been privately owned and only recently became public were excluded. Moreover, we study interlocking directorates among only the twenty-six firms. We ignore a host of board linkages to other Minnesota firms, to firms in other parts of the country, and to organizations outside of business.[4] Since research has shown that firms in the same geographical area tend to interlock with one another and not with other firms in other regions (Levine, 1972; Allen, 1978; and Sonquist and Koenig, 1975), it is reasonable to assume that our Minnesota network is a meaningful and relatively self-contained subsystem of a larger network of corporate interlocks. The interlocks among our twenty-six firms constitute approximately 14% of the total

---

[2] The twenty-six firms included in our study are: American Hoist and Derrick, Apache, Bemis, Control Data, DeLuxe Check Printers, Economics Labs, General Mills, George Hormel, Green Giant, Honeywell, International Multifoods, Jostens, Minnesota Mining and Manufacturing, Munsingwear, Pillsbury, Toro, First Bank System, Northwest BANCO, Dayton-Hudson, Gamble-Skogmo, Burlington-Northern, Northwest Orient, Soo Line, NSP, Minnesota Mutual Life Insurance, and Northwestern National Life Insurance. Despite the fact that American Crystal and Land O'Lakes were also on the *Fortune* lists for each year, they were not included in our study because American Crystal only moved to Minnesota in 1974 and Land O'Lakes is a cooperative whose board is made up exclusively of its own members. Since First Bank System and Northwest BANCO are both bank holding companies, any board ties between their respective subsidiaries—1st National of Minneapolis, 1st National Bank of St. Paul, Northwestern National Bank of Minneapolis, and Northwestern National Bank of St. Paul—and other firms on our list were recorded as ties to the two bank holding companies.

[3] We restricted our study to firms which were on our lists for all ten years because of the lack of stochastic models that allow nodes to enter and leave the network. This is indeed a problem. A more thorough analysis of these data could concentrate on year-to-year transitions, and not pool over the ten year period, thus maximizing the use of these data. However, since the emphasis here is on parameter estimation to verify our four hypotheses, pooling makes more accurate estimates (see Runger and Wasserman, 1980).

[4] Looking at all the board members of all twenty-six firms in 1969 and 1978, we find that nonprofit organizations were better represented on these boards in 1978 than in 1969. The percentage of inside directors decreased from 33.8 to 28.2, the percentage of directors from other firms in the network decreased from 14.1 to 13.1, and the percentage of directors from other business corporations decreased from 36.7 to 34.0. (The total number of board slots remained about the same over the decade: 370 in 1969 and 373 in 1978.) In contrast, the percentage of board members from nonprofit organizations increased from 4.1 to 7.2.

number of board members that the twenty-six firms employ. Moreover, we ignore characteristics of individual corporations, e.g., conditions in the various environments of the twenty-six firms, firms' individual successes and failures, and their respective industries and trading partners. We also disregard changes in firms' policies regarding board composition such as a conscious effort to increase minority or nonprofit organization representation.

Stochastic Models to Study Macrostructural Change

Recently, Wasserman (1978, 1979, 1980) and Holland and Leinhardt (1977a, 1977b) developed a class of models that depend on parameters that describe aspects of change in social networks. Most of these models are continuous-time Markov chains (see Runger and Wasserman, 1980), but we use discrete-time Markov chains (see Karlin and Taylor, 1975: Ch. 2). Discrete-time versions of these continuous-time models avoid the mathematical difficulties inherent in estimation of the structural parameters. They appear appropriate for a corporate interlock network since changes in a board of directors occur usually only once a year.

Let $\mathbf{X}_t = (X_{ij;t})$ be the adjacency matrix representing the network at time t, where

$$X_{ij;t} = \begin{cases} 1 & \text{if firm j is represented on the board of firm i at time t;} \\ 0 & \text{otherwise.} \end{cases}$$

We set t = 1 for 1969, t = 2 for 1970, . . . and t = 10 for 1978.

Reciprocity. Define $D_{ij;t} = (X_{ij;t}, X_{ji;t})$ as a variable representing the dyadic linkage between corporations i and j. A mutual (M) relationship between two corporations is represented as (1,1); a null (N) relationship as (0,0); and an asymmetric (A) as either (1,0) or (0,1). We postulate that $D_{ij;t}$ is a discrete-time Markov chain on the four states: (0,0), (0,1) and (1,0), (1,1). We also assume that $D_{ij;t}$ is statistically independent of all other dyads.[5]

To determine if there is an increased probability of a linkage (in our case, a new board interlock) developing between organizations i and j when a linkage already exists from j to i and if there is a decreased probability of a linkage between i and j disappearing when a linkage from j to i is present, we postulate the following probability transition function:

$$p_{ij;k\,l}^{t,t+1} = P\{X_{ij;t+1} = l \mid D_{ij;t} = k\}$$

where $p_{ij;k\,l}^{t,t+1}$ is the probability that the link X_{ij} is either 0 or 1 at time t+1 given that the dyad is in state k, one of the four dyad states, at time t. We assume that $p_{ij;k\,l}^{t,t+1}$ does not depend on the particular pair (i,j) so that $p_{ij;k\,l}^{t,t+1} = p_{k\,l}^{t,t+1}$ and that these conditional probabilities may be arranged in a 3 × 3 probability transition matrix $\mathbf{P}_R^{t,t+1}$ (Table 1). Because of our stipulations, the theoretical probability transition matrix, $\mathbf{P}_R^{t,t+1}$, depends solely on four parameters, λ_0, λ_1, μ_0 and μ_1. Briefly, λ_0 is the probability that a linkage will form given that no reciprocated arc existed; $\lambda_0 + \mu_0$ is the probability that a linkage will form given that there is a reciprocated arc; λ_1 is the probability that a linkage will disappear given that no reciprocated arc exists; and $\lambda_1 + \mu_1$ is the probability that a linkage will disappear given that a reciprocated arc existed. (See Runger and Wasserman,

Table 1. Theoretical Probability Transition Matrix for the Reciprocity Model

		N	A	M
	N	$1-2\lambda_0$	$2\lambda_0$	0
$\mathbf{P}_R^{t,t+1}$:	A	λ_1	$1-(\lambda_0+\lambda_1+\mu_0)$	$\lambda_0+\mu_0$
	M	0	$2(\lambda_1+\mu_1)$	$1-2(\lambda_1+\mu_1)$

[5] To assume that $D_{ij;t}$ is statistically independent of all other dyads is questionable. One of the attractions of network analysis is that it takes a holistic or systemic approach to group phenomena. Many network analysts warn that the full implications of the relationship between actor i and j cannot be understood without knowing something about these actors' relations to all other actors k and the relations among all k as well (see Newcomb, 1961, and Emerson, 1962).

1980, for a more thorough discussion of these parameters.)

Since there are twenty-six organizations in our networks, there are $\binom{26}{2} = 325$ dyads for each year. We summarize the dyad transitions in a 3×3 matrix for each pair of years. Table 2 presents the dyad transitions for the nine pairs of consecutive years 1969–1970, . . . , 1977–1978.

Popularity. To know if the probability of establishing another linkage with an actor in the network is a linear function of current popularity in the network (i.e., the number of boards on which a firm is represented) and if the probability of terminating a linkage with some actor in the network is also a function of popularity, we postulate the following probability transition function for the popularity model

$$p_{j;kl}^{t,t+1} = P\{I_{j;t+1} = l \mid I_{j;t} = k\}$$

where $I_{j;t} = \sum_{i=1}^{26} X_{ij;t}$, $j = 1,2,\ldots,26$, is the in-degree process for firm j. Obviously, $I_{j;t}$ can take on any integer value between 0 and 25, so that $I_{j;t}$ has state space $\{0,1,\ldots,25\}$. We further stipulate that $I_{j;t}$ is a Markov chain, and statistically independent of the other in-degrees. $I_{j;t}$ simply represents the number of firms on which corporation j is represented at time t. We also postulate that this function does not depend on the specific corporation j, so that $p_{j;kl}^{t,t+1} = p_{kl}^{t,t+1}$.

We stipulate that the transition probabilities depend on four parameters: λ_0 and λ_1, which measure the overall rate of change, and π_0 and π_1, which measure the effect of increased popularity on in-degree changes. These parameters can be summarized as follows:

$$P\{I_{j;t+1} = i+1 \mid I_{j;t} = i\} = \lambda_0 + i\pi_0$$

$$P\{I_{j;t+1} = i-1 \mid I_{j;t} = i\} = \lambda_1 + i\pi_1$$

for all j, and $i = 0,1,\ldots,25$. The probability transition matrix for the in-degree process, which mathematically is a linear birth and death process, is a Jacobi matrix with positive diagonal, sub-diagonal, and super-diagonal, and zeros elsewhere. The probability that the in-degree $I_{j;t}$ increases by one is a linear function of $I_{j;t}$, just as the probability that the in-degree decreases by one is also a linear function of $I_{j;t}$. Thus to answer the substantive question of whether popular firms become more (or less) popular over time, we need to estimate π_0 and π_1 in the equations above. Further discussion of these parameters is found in Wasserman (1980).

Table 3 gives the pooled in-degree transitions from year to year for our twenty-six firms. The entries are the number of firms with in-degree i at year t and in-degree j at year t+1, where t = 1969, 1970, . . ., or 1977.

Analyzing Macrostructural Change

Our first hypothesis discusses whether the probability that reciprocated board linkages become asymmetric is greater than the probability that nonreciprocated linkages become symmetric. To estimate the appropriate parameters from the reciprocity model to test this hypothesis, we compare the theoretical probability transition matrix (presented in Table 1) with the empirical probability transition matrix (constructed from the nine matrices presented in Table 2). The empirical probability transition matrix, $\hat{\mathbf{P}}_R$, is formed by pooling (or "adding over") the nine year-to-year transition matrices and then standardizing by dividing by row sums. Pooling is justified by the assumption that a stationary, first-order Markov chain is operating. In brief, $\hat{\mathbf{P}}_R$ is "fitted" to $\mathbf{P}_R$ using least squares, minimizing the distance between the observed and expected matrices. The estimation procedure is discussed in more detail in Wasserman (1980).

The estimated parameters of the reciprocity model to test hypothesis 1 are $\hat{\lambda}_0 = .009$, $\hat{\lambda}_1 = .137$, $\hat{\mu}_0 = .025$, $\hat{\mu}_1 = -.066$, $\hat{\lambda}_0 + \hat{\mu}_0 = .034$, $\hat{\lambda}_1 + \hat{\mu}_1 = .071$. The parameter sums, $\hat{\lambda}_0 + \hat{\mu}_0$ and $\hat{\lambda}_1 + \hat{\mu}_1$, are the estimated probabilities of a linkage between actors i and j forming or disappearing given the condition that j is already linked to i. Since $\hat{\lambda}_1 + \hat{\mu}_1 = .071 > \hat{\lambda}_0 + \hat{\mu}_0 = .034$, we confirm our first hypothesis. The probability that reciprocated board linkages become nonreciprocated is greater than the probability that nonrecip-

Table 2. Dyadic Transitions for Every Pair of Consecutive Years: 1969–1978

		1970			
		N	A	M	
1969	N	278	3	0	281
	A	2	33	1	36
	M	0	0	8	8
		280	36	9	

		1971			
		N	A	M	
1970	N	273	7	0	280
	A	3	31	2	36
	M	0	0	9	9
		276	38	11	

		1972			
		N	A	M	
1971	N	275	1	0	276
	A	6	32	0	38
	M	0	1	10	11
		281	34	10	

		1973			
		N	A	M	
1972	N	273	8	0	281
	A	3	30	1	34
	M	0	0	10	10
		276	38	11	

		1974			
		N	A	M	
1973	N	274	2	0	276
	A	5	32	1	38
	M	0	2	9	11
		279	36	10	

		1975			
		N	A	M	
1974	N	273	6	0	279
	A	8	28	0	36
	M	0	2	8	10
		281	36	8	

		1976			
		N	A	M	
1975	N	277	4	0	281
	A	3	30	3	36
	M	0	2	6	8
		280	36	9	

		1977			
		N	A	M	
1976	N	274	6	0	280
	A	5	29	2	36
	M	0	3	6	9
		279	38	8	

		1978			
		N	A	M	
1977	N	273	6	0	279
	A	10	27	1	38
	M	0	2	6	8
		283	35	7	

Table 3. In-Degree Transitions from Year t to Year t+1, for t = 1969, 1970, . . . , 1977

		Year t+1									
		0	1	2	3	4	5	6	7	8	Total
Year t	0	40	9	1							50
	1	9	62	5							76
	2	1	6	23	5	1					36
	3			1	8	4					13
	4			2	3	15	3				23
	5			1		1	9	2	2		15
	6						3	5	1		9
	7						3	1	3	1	8
	8								2	2	4

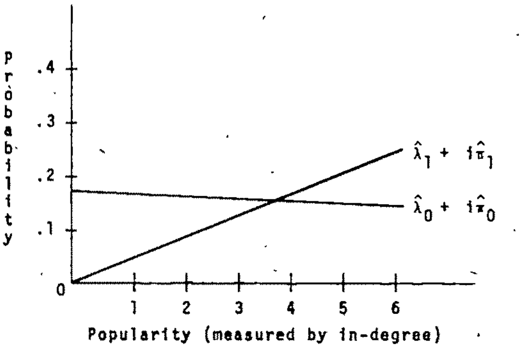


Figure 1. Popularity Model Transition Rates
NOTE: $\hat{\lambda}_0 + i\hat{\lambda}_0 = .164 - .004i$
$\hat{\lambda}_1 + i\hat{\pi}_1 = .041i$

rocated board linkages become reciprocated. Thus, there is some evidence that corporations do try to economize on information transaction costs when forming board linkages and strive to form only single-stranded links among themselves.

The second and third hypotheses discuss whether firms that are represented on several boards at time t tend to be recruited to new boards at t+1 and firms that are represented on several boards at time t tend to leave boards at t+1. To estimate the appropriate parameters from the popularity model for these hypotheses, we compare the theoretical probability transition matrix, $\mathbf{P}_p$, to the empirical probability transition matrix, $\hat{\mathbf{P}}_p$ (presented in Table 3). Again, least squares is used to measure the fit between expected and actual values.

The estimated parameters of the popularity model to test Hypotheses 2 and 3 are $\hat{\lambda}_0 = .164$, $\hat{\lambda}_1 = .000$, $\hat{\pi}_0 = -.004$, $\hat{\pi}_1 = .041$, $\hat{\lambda}_0 + i\hat{\pi}_0 = .164 - .004i$, and $\hat{\lambda}_1 + i\hat{\pi}_1 = .041i$. The functions, $\hat{\lambda}_0 + i\hat{\pi}_0$ and $\hat{\lambda}_1 + i\hat{\pi}_1$, take into account the relative impact of actors' in-degrees (or popularity) on their likelihood of going on or off boards (see Figure 1). Since the parameters are really combined simple linear functions, the graphs of these functions are useful in interpreting our findings. In this model $\hat{\pi}_0$ and $\hat{\pi}_1$ are rates of change for in-degrees; $\hat{\lambda}_0$ and $\hat{\lambda}_1$ are the y-intercepts of the equation.

To evaluate our hypotheses, we examine $\hat{\pi}_0$ and $\hat{\pi}_1$, the slopes for the functions graphed in Figure 1. The effect of a firm's in-degree (or popularity) at t on its joining a new board at t + 1 is reflected in $\hat{\pi}_0$. Contrary to our expectations, the slope is practically zero ($\hat{\pi}_0 = -.004$). The effect of a firm's in-degree (or popularity) at t on its leaving a board at t + 1 is reflected in $\hat{\pi}_1$. Here the slope is positive and in the hypothesized direction ($\hat{\pi}_2 = .041$). We had expected that more popular firms would be more likely to join new boards over time, but we found that popularity had no effect on recruitment. However, as we expected, more popular firms were more likely to leave boards than less popular firms. These findings together are somewhat puzzling. Certainly, the stratifying processes we expected are not to be found. Being represented on more boards does not make a firm especially attractive; on the contrary, it appears to hasten its departure from other boards. Perhaps firms are more interested in recruiting individuals than firms and thus ignore the number of other boards a firm is represented on. Or perhaps corporations purposely seek broader representation. Needless to say, these issues cannot be resolved here.

The fourth hypothesis is different from the previous three in that it calls for a partitioning of the dyadic transitions. It suggests that industrial, retail, transportation, and utility companies are more likely to establish new interlocks with commercial banks and insurance companies than with other industrial, retail, transportation, or utility companies. We first analyze only the empirical transition matrices that depict linkages between nonfinancial and financial firms and then analyze the matrices that depict linkages among non-

financial corporations only. The estimation procedure for the parameters is the same as that used in estimating parameters for the reciprocity model for the undifferentiated actors.

The relevant parameter estimates are $\hat{\lambda}_{0NN} = .077$, $\hat{\lambda}_{1NN} = .164$, $\hat{\lambda}_{0NF} = .013$, $\hat{\lambda}_{1NF} = .113$. The parameter, $\hat{\lambda}_{0NN}$, is the probability that nonreciprocated board linkages are formed among nonfinancial firms at t+1, given that no ties had existed at t; and $\hat{\lambda}_{0NF}$ is the probability that nonreciprocated board linkages are formed between nonfinancial and financial firms at t+1, again given that no ties had existed at t. Since $\hat{\lambda}_{0NF} = .013 > \hat{\lambda}_{0NN} = .007$, we see that there is a tendency for nonfinancial firms to link up with financials rather than with other nonfinancials.

The parameter, $\hat{\lambda}_{1NN}$, is the probability that nonfinancial board linkages with other nonfinancials, which had existed at t, are broken by t+1. $\hat{\lambda}_{1NF}$ is the probability that linkages between financial and nonfinancial firms at t are broken by t+1. Since $\hat{\lambda}_{1NN} = .164 > \hat{\lambda}_{1NF} = .113$ it appears that nonfinancials are more likely to terminate linkages with other nonfinancials than with financial organizations. Both these findings, of course, are consistent with our hypothesis and are based on the argument that board linkages between nonfinancial and financial institutions have increased over the past decade.

DISCUSSION

We have derived and tested four descriptive hypotheses taken from the managerialist and finance capitalist literatures on corporate interlocks to show how stochastic models can be used to formulate hypotheses concerning changes in corporate interlock networks. We refrained from formal hypothesis tests to judge the statistical significance of our findings, primarily because such theory is not fully developed, although there are exploratory methods that can be used to judge whether small differences between parameters are meaningful (see Wasserman, 1980, for relevant discussions). The Markov chains used here to model network processes have many additional properties, such as existence of stationary distributions and geometric waiting-time distributions. A more thorough analysis of this network should certainly study these properties.

Our nodal and dyadic models may also be insufficient because they cannot capture important tendencies toward transitivity. Such tendencies can only be verified with triadic models. To determine whether we were missing some of these tendencies by not considering a triad process, we computed the value of Holland and Leinhardt's (1975) measure for intransitivity τ (see Table 4). This standardized statistic varied between -1.2 for 1977 to 1.4 for 1972. All but two years (1977 and 1978) had positive measures, indicating slightly more transitivity than expected. Because the statistic is small, we are confident that our parameters are little affected by higher-order biases.

Substantively, our findings suggest that norms of reciprocity are not operative with respect to corporate interlocks (Gouldner, 1960), as they are in other types of interorganizational relationships. If one firm is on the board of another, there does not appear to be any pressure for the first firm to reciprocate. Board slots do not appear to be token gifts that firms exchange. In fact, when a board linkage is reciprocal, firms try to make the linkage a single bond. If board members are an instrumental communication link between two firms, this should not be too surprising.

We were surprised, however, that we did not find stratification processes at work in this corporate interlock network. Firms that were privy to more information or control by having representation on many boards were not likely to become

Table 4. Values of τ for Intransitivity

Year	τ
1969	0.691
1970	0.044
1971	1.081
1972	1.432
1973	0.686
1974	0.275
1975	1.215
1976	0.768
1977	−1.239
1978	−0.817

privy to still more information/control over time. In fact, firms that had less representation on other boards were just as likely to be recruited to new boards as firms with greater representation. If an elite exists in this network, it is not extending its control. On the contrary, it may be losing its control, for we also found that those firms that had greater representation were more likely to leave boards.

Finally, consistent with our expectations, nonfinancial institutions seem to prefer board interlocks with local financial institutions over ties with other nonfinancial institutions. This preference may simply be a function of the persistence of local financial institutions in recruiting representatives of prestigious multinationals to their boards. Nevertheless, the increasing interest in local financial institutions is noteworthy, given the alleged increased centralization in the economy. Although our data are too limited to draw any generalizations, our findings support the position that local financial institutions may be becoming an important organizing force in local business communities.

In general, our substantive results support both the finance capitalist and managerialist positions. It was not our intent to devise a critical test by pitting one theory against the other. We believe that such a test would be very difficult to construct using our models. Rather, our goal was to show how a mathematical model devised to study dynamic processes in a social network could be modified so as to entertain and test propositions drawn from different theoretical perspectives.

REFERENCES

Allen, Michael
1974 "The structure of interorganizational elite cooptation: interlocking corporate directorates." American Sociological Review 39: 393–406.
1978 "Economic interest groups and the corporate elite." Social Science Quarterly 58: 597–615.

Berkowitz, Stephen, Peter Carrington, Yehuda Kotowitz, and Leonard Waverman
1979 "The determination of enterprise groupings through combined ownership and directorship ties." Social Networks 1:391–413.

Burt, Ronald
1979 "A structural theory of interlocking corporate directorates." Social Networks. 1: 415–35.

Dooley, Peter
1969 "The interlocking directorate." American Economic Review 59:314–23.

Emerson, Richard
1962 "Power dependence relations." American Sociological Review 27:31–41.

Fennema, Meindert and Huibert Schijf
1979 "Analyzing interlocking directorates: theory and methods." Social Networks 1: 297–332.

Fitch, Robert and Mary Oppenheimer
1970 "Who rules the corporations? Parts 1, 2, and 3." Socialist Revolution 1:73–108, 61–114, 33–94.

Gouldner, Alvin
1960 "The norm of reciprocity." American Socioloical Review 25:161–77.

Herman, Edward
1973 "Do bankers control corporations?" Monthly Review 25:12–29.

Holland, Paul and Samuel Leinhardt
1975 "Local structure in social networks." Pp. 1–45 in David Heise (ed.) Sociological Methodology 1976. San Francisco: Jossey-Bass.
1977a "A dynamic model for social networks." Journal of Mathematical Sociology 5:5–20.
1977b "Social structure as a network process." Zeitschrift fur Soziologie 6:386–402.

Karlin, Samuel and Howard Taylor
1975 A First Course in Stochastic Processes, 2nd Edition. New York: Academic Press.

Levine, Joel
1972 "The sphere of influence." American Sociological Review 37:14–27.

Mace, Myles
1971 Directors: Myth and Reality. Boston: Graduate School of Business Administration, Harvard University.

Mariolis, Peter
1975 "Interlocking directorates and control of corporations: the theory of bank control." Social Science Quarterly 56:425–39.

Mintz, Beth and Michael Schwartz
1977 "The structure of power in American business." Paper presented at the Annual Meetings of the American Political Science Association, Washington, D.C., September.

Newcomb, Theodore
1961 The Acquaintance Process. New York: Holt, Rinehart, and Winston.

O'Connor, James
1968 "Finance capital or corporate capital?" Monthly Review 20:30–5.

Perlo, Victor
1957 The Empire of High Finance. New York: International Publishers.

Pfeffer, Jeffrey
1972 "Size and composition of corporate boards of directors: the organization and its envi-

ronment." Administrative Science Quarterly 17:218–28.
Runger, George and Stanley Wasserman
1980 "Longitudinal analysis of friendship networks." Social Networks 2:143–54.
Sonquist, John A. and Thomas Koenig
1975 "Interlocking directorates in the top U.S. corporations: a graph theory approach." Insurgent Sociologist 5:196–229.
Warner, W. Lloyd and Dareb B. Unwalla
1967 "The system of interlocking directorates." Pp. 121–57 in W. L. Warner, D. B. Unwalla, and J. Trim (eds.), The Emergent American Society: Large-scale Organizations. New Haven: Yale University Press.
Wasserman, Stanley
1978 "Models for binary directed graphs and their applications." Advances in Applied Probability 10:803–18.
1979 "A stochastic model for directed graphs with transition rates determined by reciprocity." Pp. 392–412 in Karl Schuessler (ed.), Sociological Methodology, 1980. San Francisco: Jossey-Bass.
1980 "Analyzing social networks as stochastic processes." Journal of the American Statistical Association 75:280–94.

COMMENTS

RECONFOUNDING THE CONFLUENCE MODEL: IS THE RELATIONSHIP OF SIBSHIP SIZE AND BIRTH-ORDER TO INTELLIGENCE AN ARTIFACT?*

(COMMENT ON STEELMAN AND MERCY, *ASR*, AUGUST 1980)

ALFRED C. HEXTER
Department of Health Services, State of California

Specialization in science can lead to situations in which scholars in one field are totally unaware of pertinent developments in another field. Something of the sort may be happening in studies which relate sibship size and birth order to intelligence.

Steelman and Mercy (1980) cite 27 papers which discuss factors which might contribute to the relationship. Factors cited include maternal fatigue, deficiencies in parenting, marital disruption, and numerous others. Notably missing, however, is condition of the child at time of birth.

It is well known to those doing research in infant health that maternal age and parity (number of previous children) are factors which affect outcome of birth, as measured by infant mortality or by birthweight (Eisner et al., 1979). Infants born to very young mothers are at especially high risk, but older maternal ages, especially mothers 35 years and older, and higher parities are also at risk. Other factors, such as interpregnancy interval (child spacing), are also important.

In recent years there has been growing recognition that mortality is not the only risk. An infant who is born with low birthweight, or whose birth was marred by complications of pregnancy or delivery, may suffer CNS involvement which ranges from severe retardation to minimal effects whose significance is not clear (Field, et al., 1979; Fisch et al., 1975; Fitzhardinge et al., 1976; Nelson and Ellenberg, 1979; Parmalee et al., 1970; Pomerance et al., 1978; Ramey et al., 1978; Salamy et al., 1980).

Another factor related to maternal age (and hence birth order) is genetic damage. The best known example is Down's syndrome, where advanced maternal age is a major risk factor. However, there is reason to believe that other, less obvious, impairments may also be related to maternal age, and some of these may affect intelligence.

I would suggest that if Steelman and Mercy still have access to the data they check the birthweights recorded on the birth certificates. The 3,428 individuals can be ranked in order by birthweight, or classed as very low, low, or normal birthweight (<1500 g, 1500–2499 g, 2500+ g). Standard methods are then available for checking whether the different groups are homogeneous with respect to birthweight rankings. Similar tests can be applied for maternal age at time of birth; and other tests are available for comparing groups with respect to occurrence of complications of pregnancy, labor, and delivery. If the groups are homogeneous with respect to birthweight and the other birth-related variables, then confounding effects from conditions present at birth may be reasonably considered to be absent. If not, then the interpretations Steelman and Mercy have presented will need to be reconsidered.

* Direct all correspondence to Alfred C. Hexter, Maternal and Infant Health Section, Department of Health Services, 2151 Berkeley Way, Annex 4, Room 500, Berkeley, CA 94704.

REFERENCES

Eisner, V., J. V. Brazie, M. W. Pratt, and A. C. Hexter
1979 "The risk of low birthweight." American Journal of Public Health 69:887–93.

Field, T. M., A. M. Sostek, S. Goldberg, and H. H. Shuman (eds.)
1979 Infants Born at Risk. New York: Spectrum Publications, Medical and Scientific Books.

Fisch, R. O., M. K. Bilek, L. D. Miller, and R. R. Engel
1975 "Physical and mental status at 4 years of age of survivors of the respiratory distress syndrome." Journal of Pediatrics 86:497–503.

Fitzhardinge, P. M., K. Pape, M. Arstikaitis, M. Boyle, S. Ashby, A. Rowley, C. Netley, and P. R. Swyer
1976 "Mechanical ventilation of infants of less than 1501 gm birth weight: health, growth, and neurologic sequelae." Journal of Pediatrics 88:531–41.

Nelson, K. B. and J. H. Ellenberg
1979 "Neonatal signs as predictors of cerebral palsy." Pediatrics 64:225–32.

Parmalee, A. H., A. Minkowski, S.S.-A. Dargassies, C. Dreyfus-Brisac, I. Lezine, J. Berges, G. Chervin, and E. Stern
1970 "Neurologic evaluation of the premature infant: a follow-up study." Biologica Neonates 15:65–78.

Pomerance, J. J., C. T. Ukrainski, T. Ukra, D. H. Henderson, A. H. Nash, and J. T. Meredith

1978 "Cost of living for infants weighing 1,000 grams or less at birth." Pediatrics 61:908–10.

Ramey, C. T., D. J. Stedman, A. Borders-Patterson, and W. Mengel
1978 "Predicting school failure from information available at birth." American Journal of Mental Deficiency 82:525–34.

Salamy, A., T. Mendelson, W. H. Tooley, and E. R. Chaplin
1980 "Differential development of brainstem potentials in healthy and high-risk infants." Science 210:553–5.

Steelman, L. C. and J. A. Mercy
1980 "Unconfounding the confluence model: a test of sibship size and birth-order effects on intelligence." American Sociological Review 45:571–82.

FAMILY STRUCTURE AND IQ: ARTIFACT OF CONDITION AT BIRTH? (REPLY TO HEXTER)*

LALA CARR STEELMAN
University of South Carolina

JAMES A. MERCY
Emory University

Along with other social scientists we heartily endorse Hexter's plea for interdisciplinary research. Specialization potentially creates situations in which scholars suffer from tunnel vision and are shielded from developments in other fields that are germane to the problem they are studying. We are surprised that this specific criticism has been leveled at our work, since from the start the plausibility of a physiological explanation of birth-order effects on intelligence is recognized (see fifth paragraph, Steelman and Mercy, 1980). Research reported in the original study is admittedly only a beginning; but, while other factors such as maternal age and birthweight could have been controlled, the indicators used in the analysis reflected the theoretical perspectives being tested. Hexter suggests that the data analyzed in the paper should be reexamined in the light of several indicators of the condition of the child at the time of its birth. We have done so, despite some misgivings about Hexter's review.

Hexter's title says explicitly that the relationships of sibling size and birth order to intelligence might be artifactual. A careful reading of our article reveals unequivocally that sibship size has a significant impact on IQ performance while birth order is unrelated to IQ. Perhaps Hexter has confused family size and birth order.

According to Hexter, we need a reanalysis to ascertain whether the significant effects of family structure on IQ performance are artifacts of associated differences among children at the time of their births. We did not and do not believe that sufficient data are available to provide a sound test of his alternative explanation. We particularly question the use of indicators such as maternal age and birthweight to show the child's condition at birth. Such factors are themselves confounded with other variables, notably race and socioeconomic status (Record et al., 1969; Caputo and Mandell, 1970). They also are not direct indicators of CNS (central nervous system) involvement.

Hexter's attempt to explain the effect of birth order as a function of condition at birth is a misfire in the case of our study, but might apply to other studies in which a significant effect of birth order emerges. We, however, doubt its relevance. Epidemiologically rare events such as Down's syndrome, although they are more common to children of higher parities and those born to older mothers, would probably not significantly alter the shape of distributions generated from large population surveys. This is especially pertinent to our study since the sample evaluated consists of *noninstitutionalized* children in which the expected number of cases with severe mental impairment is few. Hexter, however, states that there may also be subtle effects of condition at birth on IQ, not just severe retardation outcomes. While maternal age may be a useful indicator to estimate the probability of outcomes such as Down's syndrome, we have seen no strong evidence that it predicts other outcomes such as IQ performance. Age of the mother is related to her emotional maturity, childrearing skill, and other life experiences that may offset or even reverse the slight biological disadvantages associated with being born to older mothers. In addition, the information taken from the crossclassifications between maternal age and child's IQ and birthweight and IQ is equivocal. Inconsistencies are noted across cultures (Omran et al., 1976). Examination of the one-way distributions among these variables provides only a rough image of the impact of the child's condition at birth, since birthweight and maternal age are related to other variables that ought to be taken into account (Record et al., 1969; Caputo and Mandell, 1970).

In reexamining our data with Hexter's reappraisal in mind, we first performed some simple crosstabulations to see whether the indi-

* Direct all correspondence to Lala Carr Steelman, Department of Sociology, University of South Carolina, Columbia, South Carolina 29208.

We are grateful to Carl W. Tyler Jr. for his remarks on an earlier draft of this reply and to Brent Blumenstein for his suggestions.

Table 1. Unstandardized Regression Coefficients and Their Standard Errors (In Parentheses) Predicting IQ Performance

Variable	Full Model (1)	Final Model Including Pregnancy Items (2)	Final Model Excluding Pregnancy Items (3)
Below-Poverty Status (N = 424)			
Income	−2.1414 (.5539)	2.0288 (.5439)	2.3694 (.5499)
Sibship Size	−.9159 (.4574)	−1.2955 (.2943)	−1.3635 (.2990)
Birth Order	−.7712 (.7151)		
Mother's Working Status	−.6557 (1.6002)		
Mother's Age	−1.8339 (2.5539)	.4170 (.1192)	
Birthweight (grams)	−.0235 (.0148)	−.0228 (.0129)	
Pregnancy Length (weeks)	−2.2420 (1.8464)	−.7565 (1.0103)	
Mother's Age × Birthweight	.0000 (.0002)		
Mother's Age × Pregnancy Length	.0592 (.0654)		
Birthweight × Pregnancy Length	.0006 (.0003)	.0006 (.0003)	
Intercept	163.7271	107.1238	91.4975
Adjusted R^2	.1306	.1343	.0934
Above-Poverty Status (N = 2428)			
Income	3.0392 (.2302)	3.0413 (.2298)	3.1662 (.2288)
Sibship Size	−.9947 (.2186)	−1.1843 (.1574)	−1.1312 (.1578)
Birth Order	−.3978 (.3117)		
Mother's Working Status	1.3081 (.5600)	1.4044 (.5552)	1.6129 (.5572)
Mother's Age	−.1339 (1.3485)	.1941 (.0480)	
Birthweight (grams)	.0222 (.0087)	.0234 (.0084)	
Pregnancy Length (weeks)	1.5385 (1.0927)	1.6834 (.5952)	
Mother's Age × Birthweight	.0000 (.0001)		
Mother's Age × Pregnancy Length	.0055 (.0366)		
Birthweight × Pregnancy Length	−.0006 (.0002)	−.0006 (.0002)	
Intercept	18.0551	9.1299	83.8706
Adjusted R^2	.1069	.1073	.0951

Total N = 2852

cators listed by Hexter—maternal age, birthweight, and complications during delivery—are associated with IQ. Length of gestation is also included as an indicator of prematurity. Although Hexter does not recommend explicitly the inclusion of gestation length, others underscore its connection to birthweight and to prematurity (Caputo and Mandell, 1970).

Groups are divided along the poverty line and the sample is restricted to white children from intact families.[1] Under the poverty line, IQ increases by mother's age up until ages 35 and over. Above the poverty line, IQ increases with mother's age except for mothers over 40. The relationship between birthweight and IQ is anomalous. The breakdown advocated by Hexter is used (<1500 grams, 1500–2499 grams, 2500+ grams). IQ is lower only for children in the middle category, 1500–2499 grams, contrary to expectation, across both income groups. IQ rises uniformly across all categories of gestation length for children in the below-poverty group. Among children in the above-poverty group, IQ is lower for children born early in the 26–36 week category but is not significantly different among higher gestation length categories (37–41, 41+). Numerous indicators of childbirth trauma are assessed. None show significant results.[2]

Although the univariate distributions suggest some important connections between the so-called condition-at-birth indicators and IQ performance, the question of whether these indicators could actually explain the impact of family size found in the original study requires regression analysis. Column 1 in Table 1 presents the unstandardized regression coefficients yielded when predicting IQ from family size, income, mother's age, mother's working status, birthweight, birth order, length of pregnancy, and three interaction terms.[3]

[1] See Steelman and Mercy (1980) for further details on how the cases are selected.

[2] Items available on childbirth difficulties in the data base are rough. As an illustration, one question that we looked at involved simple yes/no responses to an item asking if there had been any difficulties at birth. In regard to intellectual development, specific questions concerned with hypoxia (oxygen deficiency) or other brain damage might be more pertinent to the issue but are unavailable.

[3] Preliminary investigation included squared terms of mother's age and birthweight to hunt for possible curvilinear relationships. However, the squared

The number of cases is less than in the original 3,428 because of missing information on the newly introduced indicators. Columns 2 and 3 are of central interest since they compare the results obtained using the original study variables, reestimated given the reduced number of cases, with the results from the reanalysis. Table 1 indicates clearly that the effects of family size and income are slightly attenuated with the inclusion of the condition-of-birth indicators. A test of the null hypothesis that the corresponding slopes for income and sibship size are equal across the models presented in columns 2 and 3 (in both the below- and above-poverty conditions) reveals no significant differences (Dixon and Massey, 1969:209).

Nonetheless, the independent effects of the indicators of condition at birth that are significant merit discussion. Maternal age is associated positively with IQ. Apparently, mother's age is not just a biological indicator, since having an older mother seems beneficial. In addition, the interaction effect of birthweight and gestation length is significant for the above-poverty group at or beyond the .05 level but is significant for the below-poverty group only at the .10 level.[4] Upon closer inspection of the crosstabulation between these two variables and IQ, it seems that the combination of short-term pregnancies and low birthweight is associated with lower IQ. This conclusion is arrived at cautiously because there are too few cases of low-birthweight babies. Future researchers should explore these ideas with samples that provide an adequate number of cases.

No support for Hexter's assertions is found. Still, investigators should continue to search among various alternative explanations to account for the inverse effect of family size on IQ. Study of this topic ensures that researchers can make a contribution that cuts across disciplinary boundaries. Although Hexter puts forth an interesting argument, the primary conclusions from our paper stand unchallenged. The scholar is welcomed whose own work places ours into obsolescence.

REFERENCES

Caputo, Daniel V. and Wallace Mandell
1970 "Consequences of low birth weight." Developmental Psychology 3:363–83.

terms did not provide a better fit of the data than their unsquared counterparts. This may be attributable to the small number of cases that appear at the extreme ends of the distributions of these variables.

[4] The below-poverty ordinary least-squares equation was reestimated dropping out the interaction terms. With these terms excluded, gestation length and mother's age both are associated positively with IQ, whereas birthweight is unrelated to IQ.

Dixon, Wifrid J. and Frank J. Massey
1969 Introduction to Statistical Analysis. New York: McGraw-Hill.
Omran, A. R., C. C. Standley, J. E. Azar, M. Bertan, V. Guzman, V. Nahapetian, and K. A. Pisharoti
1976 Family Formation Patterns and Health. Geneva: World Health Organization.
Record, R. G., T. McKeown, and J. H. Edwards
1969 "The relation of measured intelligence to birth order and maternal age." Annals of Human Genetics 34:11–20.
Steelman, Lala Carr and James A. Mercy
1980 "Unconfounding the confluence model: a test of sibship size and birth-order effects on intelligence." American Sociological Review 45:571–82.

CAREER PATTERNS OF SCIENTISTS: A CASE FOR COMPLEMENTARY DATA*

(COMMENT ON LONG ET AL., *ASR*, OCTOBER 1979)

DARYL E. CHUBIN
ALAN L. PORTER
Georgia Institute of Technology

MARGARET E. BOECKMANN
The Urban Institute

Introduction

Fascination with "status attainment" as a fundamental dimension of social structure and institutional stratification now pervades the sociology of science (Hargens, 1978:123–236). A corresponding dependence on unobtrusive data sources has been made possible by such inventions as the *Science Citation Index*, biographical compendia such as *American Men and Women of Science*, and graduate program ratings compiled by the American Council on Education. While we, too, have utilized these sources for reconstructing scientists' career patterns (Chubin, 1974; Crowley and Chubin, 1976; Porter and Wolfle, 1975; Porter, 1977), our fascination differs from the genre of analysis reported in this journal.

For example, two recent studies of careers of biochemists (Long, 1978; Long et al., 1979) focused exclusively with unobtrusive data on prestige or status attainment and ignored other dimensions that the literature on career patterns of scientists has illuminated. In response, this comment seeks to clarify the limits of gen-

* Direct all correspondence to Daryl E. Chubin, School of Social Sciences, Georgia Institute of Technology, Atlanta, GA 30332. The data presented here were collected under NSF Grant No. SRS78-18959 awarded to Alan L. Porter, Principal Investigator. We thank the three quasi-anonymous referees for constructive comments.

eralizing the results from the study of Ph.D.s in one scientific field to other fields (which Long et al., 1979:829, deem unproblematic), and to augment the analyses of 1957–1958 and 1962–1963 Ph.D. cohorts in biochemistry with a comparative analysis of 1969–1970 Ph.D.s in six fields, including biochemistry. Through the latter, we intend to show that scientists' self-reports complement findings based solely on public-domain data. Without "subjective" data, we argue, the mechanisms that underlie career decisions and generate observed patterns of employment, productivity, and prestige remain obscure.

Complementary Data

Our data base consists of 645 responses to a six-page questionnaire mailed February 1979 to a random sample of 1969–1970 Ph.D. recipients from U.S. universities. The sample was stratified by field with an overall response rate of 70% (see Appendix). Among the areas we surveyed are the following: the role of the student's dissertation supervisor or mentor (including coauthorship with him/her prior to the student's first post-Ph.D. job), the full-time equivalent (FTE) years employed in business/industry, government, and academe; and the extent to which research has been the primary work activity.

These data are retrospective accounts of career experiences over a ten-year period and are subject to scientists' rhetoric—the omissions, juxtapositions, and embellishments of their recollections (Mulkay, 1974, 1976; Woolgar, 1976; Studer and Chubin, 1980). We are not claiming the infallibility of scientists' accounts, but only that recollections elicited in response to closed- and open-ended questions about a memorable time in the scientist's professional life can augment information gathered from archival sources. The survey data we have gathered can provide insights into processes, motives, and intentions—the behaviors which all-too-routinely are attributed to scientists in the absence of direct measurement.

Beyond Prestige: Sectors of Employment. Table 1 summarizes the cross-field comparisons on seven key variables for our 1969–1970 Ph.D. sample. Among other findings, it shows great variation by field in the extent of academic employment among the cohort. Specifically, a substantial portion of the biochemists' first post-Ph.D. decade was spent in nonacademic settings. Long (1978:892) notes the atypicality of productivity patterns among biochemists who leave academe (examined below), yet is willing to generalize from academic biochemists to all scientists.

The primary work activities of our sample vary strikingly across fields, with biochemistry an extreme. That is, research appears to be more prominent for biochemists; hence, it might be expected to relate more strongly to their professional attainment than would be the case for other scientists. A general reduction in research and teaching with a commensurate increase in administration is also evident over the decade, but biochemistry is still the extreme.

As Zuckerman and Merton (1972) suggest, transitions out of research activity adversely affect productivity. Too often, however, productivity studies such as Long et al.'s (1979) tend to link positional status with publication productivity alone, thereby neglecting other "local" factors that possibly contribute to prestige, such as teaching and administration (see Whitley, 1977). Instead of assuming that moribund research "pushes" scientists out of that activity, we must consider the "pull" of other activities and roles which themselves command prestige or its surrogates, such as high remuneration. This may especially be the case in nonacademic sectors of employment (Marsh and Stafford, 1967) where appropriate performance measures are lacking (Reskin, 1979:144).

Reinterpreting the Postdoctoral Appointment. Certainly a factor in any "prestige tradeoff" would be the market conditions faced by the new Ph.D. For example, we would expect differences in the early careers of a cohort of scientists trained prior to Sputnik and of one awarded the Ph.D. thereafter. Long et al.'s (1979) two cohorts straddle this particular reality—a "social shock" that reshaped the employment market for Ph.D. scientists throughout the 1960s—yet they are aggregated for analysis. The 1969–1970 cohort confronted a different reality when embarking on their careers: the beginnings of a shrinking, inflation-wracked economy.

We not only asked our sample whether a postdoctoral position was obtained, but also inquired as to the reason for taking it. As inferred from Table 1, the glut of Ph.D.s produced since Sputnik had begun to take its toll on the job market, particularly on physicists. Part of the strategy in coping with such a market—actually comprised of national, regional, and special-interest submarkets (Hargens, 1969; Brown, 1967)—is the acquisition of new skills to enhance one's employability, i.e., to extend one's capabilities beyond those certified by receipt of the doctorate. Market conditions, or perceptions thereof, can affect decisions that alter the course of a career. These same conditions un-

Table 1. Selected Cross-Field Comparisons Among 1969–1970 PhDs

Variable	Biochemistry	Electrical Engineering	Physics	Psychology	Sociology	Zoology	All
Total N	119	106	97	107	93	123	645
Academic Employment:							
%[a] reporting $\geq$ 50% FTE years academic[b] (mean FTE years academic)	65.5 (5.7)	29.5 (2.6)	50.5 (4.9)	57.5 (5.3)	89.2 (8.6)	90.2 (8.2)	64.3 (5.9)
Primary Work Activity: First Post-Ph.D. Job (Current Job)[c], in %[a]							
–Research, or research and teaching	90.7 (67.0)	36.5 (26.4)	54.3 (34.7)	29.9 (18.1)	25.3 (24.4)	51.6 (53.7)	49.4 (38.6)
–Teaching	3.4 (4.3)	13.5 (4.7)	27.7 (22.1)	27.1 (21.0)	53.7 (36.7)	45.1 (29.8)	29.2 (19.3)
–Administration, or administration and research or teaching	0.0 (9.6)	1.9 (25.5)	0.0 (14.7)	1.9 (20.0)	5.5 (28.9)	0.8 (11.6)	1.6 (17.9)
Postdoctoral Positions Held, in %[a]	72.3[d]	9.5	40.2	26.2	18.5	41.5	35.9
Reasons Given for Taking Postdoctoral Positions, in %[e]							
–Research experience	70.6	30.0	53.8	46.4	53.3	83.3	64.4
–Switching specialty area	14.1	10.0	2.6	28.6	33.3	8.3	13.8
–Lack of desirable permanent employment	11.8	30.0	43.6	0.0	0.0	6.3	14.7
Coauthored a Publication with Dissertation Supervisor, in %[a]	91.4	62.7	65.3	44.3	26.1	45.5	56.6
Early Career Publication[f]:							
% with no journal articles	7.6	35.8	14.4	34.6	36.6	17.9	23.9
% with 10+ journal articles	49.6	12.3	28.9	23.4	21.5	36.6	27.6
Mean (Median) articles	11.1 (9.4)	4.6, (2.1)	7.6 (4.6)	5.4 (2.4)	5.3 (3.2)	9.0 (6.6)	7.3 (4.5)
Received Important Assistance from Their Supervisor in Obtaining First Post-Ph.D. Job, in %[a,g]	37.2	15.8	31.2	20.2	22.2	26.5	25.9

[a] Percentages are based on number responding unambiguously (i.e., blanks and special situations excluded).

[b] Full-time equivalent years spent employed in academic institutions divided by the sum of FTE years academic, industry, government, and other.

[c] Two categories are excluded: "professional service" (e.g., 37.4% of the psychologists' first activity) and "development and design" (e.g., 39.6% of the electrical engineers' first dominant activity); hence percentages do not sum to 100.

[d] The great majority of the biochemistry postdoctoral appointments took place immediately after completing the doctorate; this is consistent with Long et al.'s (1979:893) 65% so engaged.

[e] "Other" category not shown; it is the difference from 100%. Note that the Ns are low in some fields (e.g., 10 postdocs in EE in all).

[f] Source: SCI *Source Index* (SSCI as well for sociologists and psychologists) corrected by reports of survey respondents.

[g] Percent noting "very important" or "important" on a 5-point scale.

derscore the social mechanisms, typically conceptualized as particularism vs. universalism, which affect not only entrance to the scientific career but also the configuration of paths taken thereafter.

Particularism vs. Universalism Reconsidered. Figure 1A attempts to simplify the key features of the argument presented by Long et al. Employing a series of regressions based on a longitudinal data base, they conclude that prestige of the first job strongly associated with prestige of the most recent prior departmental affiliation (i.e., postdoctoral appointment or Ph.D. granting institution), moderately associated with mentor's prestige, and negligibly associated with pre-Ph.D. publications. Conversely, they find the best predictor of later career publications to be pre-Ph.D. publications, while Ph.D. rating, postdoc rating, and mentor prominence are unimportant (see Table 2, part A). Long et al. conclude that particularism is at work inasmuch as first academic appointments relate to reputational factors irrelevant to later scientific productivity, but not to pre-Ph.D. publications which do relate to later productivity.

We have selected a set of variables derived from our data base to gauge the comparability between relationships observed for the 1969–1970 cohort and Long et al.'s findings. Figure 1, Part B depicts our model with the relationships observed, while Table 2, Part B, provides the numerical details of those observations. Our measures of Ph.D. rating, early publications, and prior publications are conceptually similar to Long et al.'s, although operationally somewhat different. We consider whether a postdoctoral appointment was taken rather than the rating of its prestige, and we approach mentor prestige quite differently (see notes to Table 2). Lacking prestige ratings in our Ph.D.s' employment settings, we sought an alternative to publication and citation as measures of professional stature. We were directed to salary as the most suitable variable, as based

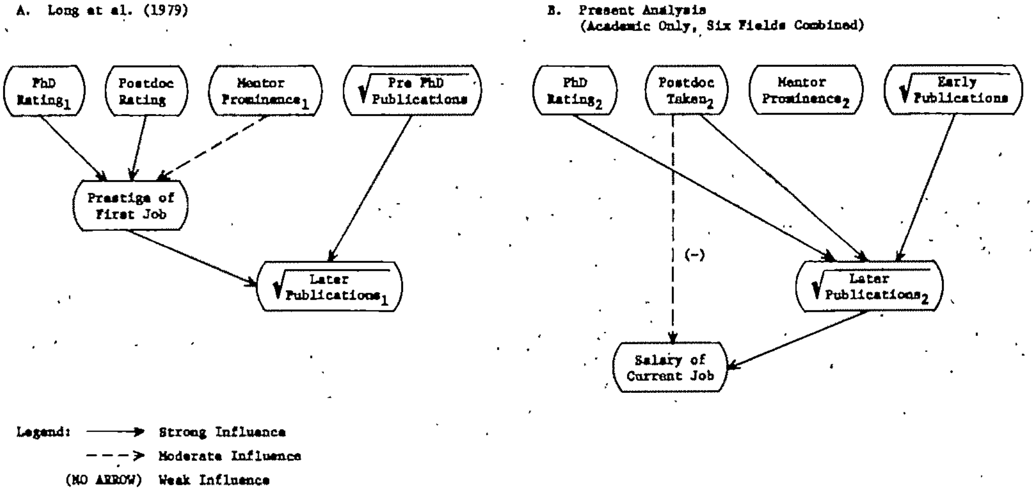


Figure 1. Simplified Models

NOTE: See Table 2 for definitions and details. Variables are not operationally equivalent in the two models. Some variables used in each analysis are omitted from A and B.

on extensive studies by the National Research Council (Harmon, 1963; Harmon, personal communication). Current salary chronologically follows later publications (but multiple panel data to uncouple these measures satisfactorily are not available). This distinguishes our model from Long et al.'s in that prestige of first job antedates later publications. We report regression coefficients predicting later publications among biochemists and for the academic scientists in our five other fields (Table 2). In Table 3, we summarize salary predictions for the two employment sectors (academic vs. nonacademic) in each of the six fields.

We suggest three interesting counterpoints to Long et al. (1) The linkage between mentor stature and first job must be challenged. Their citation-based mentor stature measure potentially confounds work of coauthors with the protege's own. Indeed, Table 1 indicates that over 90% of our biochemists coauthored with their dissertation supervisors (far more than in any of our other five fields);[1] thus, mentor stature has diminished explanatory power vis-à-vis the students' placement. Furthermore, the pre-Ph.D. publication rate may be inherently correlated with mentor stature as measured by Long et al. Our mentor prominence measure is conceptually independent and is virtually uncorrelated with pre-Ph.D. publication rate (r = .08 for all respondents). It does correlate with Ph.D. institution rating (r = .30 for all respondents, but only .14 for academic biochemists).[2]

(2) The moderate linkage between mentor prominence and first job prestige reported by Long et al. deserves reconsideration. As shown in Table 1, only about one in three of our biochemists—and this is a greater proportion than in other fields—perceived that their mentors were important in their securing a first job.[3] Context, too, must be weighed. For in-

[1] Sixty-nine percent of the biochemists related that a key reason for selecting their dissertation topic was its preference by their supervisor. The student's personal interest in the topic, its manageability, and potential intellectual contribution were considerably less influential. A similar pattern obtained only for the physicists; in all the other fields personal interest dominated faculty preference as an influence on topic choice.

[2] Notably, from our cohort's perspective, a key reason for selecting the Ph.D. program in which they were trained was the availability of financial support: over 60% of the biochemists, psychologists and physicists cited this—more than deemed important reputation of the university, department, or speciality; presence of a particular faculty member; or situational factors (responding 4 or 5 on scales from 1 = not important to 5 = very important). For the sociologists and electrical engineers, reputation of the department and the university, respectively, most influenced their choice of graduate program. Only the zoologists reported that the presence of a faculty member most affected their choice. The 63% reporting this as a vital selection criterion contrasted with the 25% of the sociologists, psychologists, and physicists for whom particular faculty were a major influence in their decision to enter a program (multiple choices were permitted).

[3] However, as one anonymous referee suggested, "those who get really good jobs can attribute it to their own talents; those who don't get good jobs can also say their mentors did not help them." Such rationalizations no doubt occur. But the accounts of the Nobel laureates interviewed by Zuckerman

Table 2. Regressions Relating Scientists' Status and Performance to Pre- and Post-PhD Career Variables

A. Long et al. (1979)—Fiscal year 1957, 1958, 1962, 1963 PhDs

Equation		PhD Rating$_1$	Postdoc Rating$_1$	Mentor Prominence$_1$
1. Prestige of biochemists' first academic job (N = 239)	b	.324**	.277**	.390*
	r	.390	.269	.342
2. $\sqrt{\text{Later publications}_1 \text{ of academic biochemists who did not change institutions}}$ (N = 134)	b	.000	.001	−.011
	r	.165	.189	.159

B. Present Analysis—1969–1970 PhDs

$\sqrt{\text{Later Publications of:}}$		PhD Rating$_2$	Postdoc Taken	Mentor Prominence$_2$
3. Academic biochemists (N = 64)	b(r)	.287 (.167)	.095 (.140)	.069 (−.016)
4. Nonacademic biochemists (N = 33)	b(r)	.275 (.253)	.197 (.332)	.043 (−.018)
5. Academic electrical engineers (N = 28)	b(r)	.993* (.255)	−.037 (.143)	.394 (−.051)
6. Academic physicists (N = 42)	b(r)	.274 (.329)	.182* (.410)	−.038 (−.245)
7. Academic psychologists (N = 50)	b(r)	−.182 (.006)	−.009 (.186)	−.024 (−.107)
8. Academic sociologists (N = 72)	b(r)	.157 (.241)	.017 (.028)	.097 (.142)
9. Academic zoologists (N = 93)	b(r)	.256 (.331)	.139**(.375)	.041 (−.211)
10. Academic, all fields (N = 349)	b(r)	.241**(.235)	.118**(.310)	.108 (−.074)
11. Nonacademic, all fields (N = 192)	b(r)	.165* (.151)	.151**(.314)	.056 (.004)

Note for Part A: Dependent variable in Equation 1 is the Roose-Andersen (1970) bioscience prestige score of the first academic position (Long et al.: 819). In Equation 2 it is the square root of standardized (Long et al.: 819) journal publication levels for the three-year period ending in the sixth year of the first job. PhD rating$_1$ = Cartter (1966) prestige of the PhD department; Postdoc Rating$_1$ = Roose-Andersen bioscience prestige of fellowship location for fellows in rated departments, 358 for others; Mentor Prominence$_1$ = square root of five-year citation counts for mentor; $\sqrt{\text{pre-PhD Publications}_1}$ = square root of standardized levels of three-year publication counts ending in the first year of the first job; Selectivity of Undergraduate Institution = scale score from Astin (1971) (see Long et al.: 819); Enrollment in Graduate Department is for 1962; $\sqrt{\text{citations to Pre-PhD Publications}}$ = square roots of standardized values of citations to publications in the three-year period ending in the first year of the first job.

Note for Part B: Dependent Variable in Equations 3–11 is the square root of articles published 1975 and later. (Pearson correlation between this SCI/SSCI measure (see Table 1) and a measure that includes journal articles plus books, book chapters, and proceedings is .98 for the combined field sample.) PhD rating$_2$ is the complete score Roose-Andersen rating of the PhD department (kindly provided by C. J. Andersen); Postdoc Taken is a yes/no item; Mentor Prominence$_2$ = 1–4 scaled perception item; $\sqrt{\text{Early Publications}}$ = square root of publications predissertation and those derived directly from the dissertation; coauthor with Mentor is a yes/no item; Mentor Aid in Securing First Job = 1–5 scaled item on the dissertation supervisor's importance therein; $\sqrt{\text{Early Citations}}$ = square root of citations in 1970–1971 to work published through 1971.

stance, in an abundant market (plentiful positions for few new doctorates), active efforts by a mentor may be less important than in the opposite situation. Such conditions are apt to be quite dynamic and field-specific (witness the bust in the engineering Ph.D. market of the early 1970s and the current boom). It is interesting to note, too, that for our cohort, neither the measure of mentor prominence nor of mentor aid in securing the first job are significant predictors of later publication productivity (Table 2) or, with the exception of zoology (where mentor aid is a negative influence), of salary (Table 3).

(3) Long et al. claim that pre-Ph.D. publications are the best available predictor, at the time of receipt of the doctorate, of anticipated later contributions (Table 2, Equation 2). Our early publication measure is significantly predictive in two fields—sociology and zoology (Equations 8 and 9). Furthermore, our regressions suggest that Ph.D. institution prestige rating is salient only for electrical engineers (Equation 5). Beyond the doctorate, the taking of a postdoctoral appointment contributes significantly to the prediction of later publications for physicists (Equation 6) and zoologists, as do early citations for psychologists (Equation 7) as well as zoologists.

Equations 2 and 3 provide the most direct comparison between our results and those of Long et al. We confirm their finding of a posi-

(1977), plus the written commentary volunteered by over one-third of the respondents to our survey (see Chubin, forthcoming), show that personal vanity and accreditation of mentors do not strictly correspond to one's own eminence or lack of career success, respectively. Rather, there are continua of credit, admiration, and hostility which the passage of time and scientists' own rhetoric dulls or intensifies. Attained status alone does not predict where on those continua a scientist will fall.

Table 2. *(continued)*

$\sqrt{\text{Pre-PhD Publications}}$	Prestige of First Job	Selectivity of Undergraduate Institution	Enrollment in Graduate Department	$\sqrt{\text{Citations to Pre-PhD Publications}}$	Intercept	R^2
−4.38	—	6.67*	−4.54**	3.69	37.8	.275
.143	—	.223	.054	.221		
.394**	.003**	.059	−.001	−.022	−.188	.211
.332	.331	.167	−.027	.243		

$\sqrt{\text{Early Publications}}$	Co-author with Mentor	Mentor Aid in Securing First Job	$\sqrt{\text{Early Citations}}$	Intercept	R^2
.239 (.158)	−.152 (.060)	.108 (.133)	.009 (.086)	.825	.102
−.050 (−.124)	.017 (.216)	−.285 (−.281)	−.016 (.107)	.603	.251
.909 (.308)	−.027 (.184)	.359 (.288)	−.462 (.150)	−4.000	.307
.293 (.323)	.153 (.382)	.035 (.398)	.117 (.343)	−.901	.394
.297 (.378)	.555 (247)	.034 (.099)	.501** (.535)	1.127	.361
.709** (.603)	.082 (.305)	.015 (.089)	.094 (.381)	−.895	.402
.348** (.394)	.100* (.269)	−.008 (.113)	.640** (.367)	−.445	.376
.376** (.383)	.092** (.306)	.066 (.195)	.113** (.271)	−.588	.279
.227** (.238)	.066* (.222)	−.053 (.061)	.180* (.264)	−.451	.189

N = number of observations included in the regression. *r* = zero-order product moment correlations with the dependent variable; *b* = unstandardized regression coefficients; ** = p ≤ .05, two-tailed t-test; * = p ≤ .10, two-tailed test. Note that several of our variables are not interval scaled, hence the regression and correlation coefficients should be taken as suggestive, not definitive. Our multiple regressions were performed using SPSS's listwise deletion option. Both stepwise regressions and pairwise deletion were run for the equation predicting $\sqrt{\text{Later Publications}}$ yielding only trivial differences in the magnitudes of coefficients.

tive association between early and later publication rates (although our b is not significant), but we find that the correlation between Ph.D. institutional rating does not wash out in the regression—the b is large (though not statistically significant). On balance, our results diverge from those of Long et al. The magnitude and significance of the coefficients predicting

Table 3. Summary of Regressions Relating Salary[a] to Career Variables, by Field and Sector

Field	Academic/ Nonacademic	R^2	Major Predictor Variables[b]	Observed Influence[c]	R^2 in Salary Accounted for by Major Predictor Variables
Biochemistry	A	.156	$\sqrt{\text{Later pubs}}$	+	.107**
	N	.234	$\sqrt{\text{Early cites}}$	−	.128*
Electrical Engineering	A	.270	Mentor coauth	−	.143
	N	.029	—		—
Physics	A	.164	Postdoc, mentor aid	−, −	.045, .038
	N	.165	PhD Rating	+	.067*
Psychology	A	.274	$\sqrt{\text{Later pubs}}$, $\sqrt{\text{Early cites}}$	+, −	.121**, .063*
	N	.346	Postdoc, mentor coauth	−, −	.156**, .125**
Sociology[d]	A	.147	$\sqrt{\text{Later pubs}}$, postdoc	+, +	.053*, .047
Zoology[d]	A	.124	Mentor coauth, Mentor prom	+, −	.054**, .046**
All fields	A	.050	$\sqrt{\text{Later pubs}}$	+	.032**
	N	.067	Postdoc	−	.034*

[a] Salary = current professional income reported in 11 categories of $4000 increments (multiplied by 1.22 for those employed on an academic year basis).

[b] Criterion for inclusion is substantial contribution of the variable to overall R^2 in the multiple regression equation.

[c] + or − denotes sign of the regression coefficient or observed influence on salary.

[d] Low N for sociology and zoology academics precludes analysis.

* = p ≤ .10

** = p of unstandardized regression coefficient of predictor variable ≤ .05

future publications of biochemists vary even for our respective subsets of academic biochemists. The present data reveal quite different predictions of future publications across five fields of science and one engineering discipline. Interestingly, only for nonacademic biochemists does the battery of predictor variables account for more variance in later publication than it does for academic productivity (see R^2 column; nonacademic equations for other five fields not shown). Thus, the field differences in early career publication observed in Table 1 have been explored, with modest success, in Table 2. As Equations 10 and 11 show for all fields combined, this success is moderately better for predicting academics' than nonacademics' later publication. Nevertheless, this is no cause for rejoicing and further cautions against aggregating fields in future analyses.

Concerning the prediction of salary, Table 3 yields mixed results. The R^2 values are generally smaller than those for the later publication equations. For two fields (biochemistry and psychology), the salary predictions for nonacademics exceed those for academics. This is reversed for electrical engineering and there is no difference in R^2 for physics. The most consistent predictor of academics' salaries—in biochemistry, psychology, and sociology—is later publication. The early citation measure is a significant negative predictor in one academic (psychology) and one nonacademic (biochemisty) field. Two of the predictor variables are uniformly negative with single exceptions—the taking of a postdoc, in sociology, and coauthoring of at least one paper with one's mentor, in zoology. Except for these heavily academic fields, lacking a postdoc experience and not publishing with one's mentor contribute to current salary, significantly so for nonacademic psychologists.

The major conclusion to derive from our salary data is that performance norms vary by sector. This cautions against generalizing from academics to nonacademics because the reward structures of their respective employment organizations differ. Publication is not necessarily valued in nonacademic settings, and a salary earned there appears to depend on variables other than those either we or Long et al. included in our models.

Implications and Prospects

Concerning academic biochemists, our results support Long et al.'s contention that early publication predicts later publication; their observation that Ph.D. institutional prestige does not predict later publication productivity, however, is not sustained. Overall, Long et al. excelled in predicting later publication productivity of institutionally immobile academic biochemists. We excelled in predicting both the later publication and the salary of nonacademic biochemists. Perhaps the chief convergence of our respective analyses is support for the influence of "pre-Ph.D." or "early" publication on later productivity (research) and achievement (salary). If we consider other fields and institutional settings beyond biochemistry and academe, we find striking differences in pre- and early post-Ph.D. career experiences. Surely self-reports and models alone will not do. So what, we may ask, are the policy and theoretical implications of these findings?

In terms of policy, one could surmise that informal networks of communication, though neutralized in principle by affirmative action, are still extensively utilized in the placing of new Ph.D.s (e.g., Reskin, 1978). Perhaps it is time for a replication of Caplow and McGee's (1958) classic study of the marketplace, with nonacademic sectors included. Bucher and Stelling's (1977: chapter 5) observational/interview study of the professional socialization process, including a sample of doctoral biochemistry students, would be instructive for fleshing out any field-specific[4] "mentor effect" (also see Krohn, 1971; Reskin, 1979) that unobtrusive methods only begin to illuminate.

As for theoretical repercussions, instead of assuming scientists' single-minded quest for status and questioning the observance of universalism in that quest, perhaps career-patterns analysts should undertake a reappraisal of *all* the norms as they bear on the unfolding scientific career. Merton (1942, 1965) articulated a scientific ethos whose component norms engender ambivalence in the scientist and are not readily distinguishable empirically (for a review, see Stehr, 1978). Thus, while the

[4] Scientific field may not even be the appropriate level of analysis. The subfield or specialty may circumscribe a network of signficant others who are used for placing new Ph.D.s in certain sectors and settings (e.g., government as opposed to medical school laboratory) (Studer and Chubin, 1980: chapter 2). Our best approximation with the present data are a series of regressions predicting "% FTE academic" by field. Again, the evidence favors field-specific interpretations. In five of the six fields, mentor variables (coauthorship with, prominence of, aid in securing first job) are the major predictors. However, the R^2 values range only from .027 (zoology) to .20 (sociology). In the latter field, mentor prominence is positively related to a career in academe; but in zoology and psychology (R^2 = .132), this is negatively related. Finally, taking a postdoc is the only significant (and positive in sign) predictor of % FTE academic in biochemistry (R^2 = .098).

prescriptive content of the norms is widely acknowledged, their descriptive content remains blurred (e.g., Barnes and Dolby, 1970; Mitroff, 1974; Mulkay, 1980). What, in short, does the operation of particularism in the placement of some Ph.D.s reflect about the hypothesized communality, organized skepticism, and disinterestedness that characterize what some (e.g., Nelkin, 1975) maintain is an academically stereotyped reward system of science?

Clearly, disparate assumptions underpin models of the scientific career (see Bourdieu, 1975; Knorr, forthcoming) and predispose analysts to certain variables, data, and interpretations. In view of our findings and those assembled by Long and his colleagues, the prospect of a continuing reappraisal of career patterns—via complementary approaches as well as data, on past cohorts as well as contemporary ones—would seem welcome indeed.

APPENDIX

A PROFILE OF THE SIX-FIELD SAMPLE

The range in response to the questionnaire extends from a low of 64.2% in physics to a high of 79.4 in zoology (76.3% in biochemistry). The number of usable responses ranges from 93 in sociology to 123 in zoology (119 in biochemistry). Comparison of our sample's characteristics (e.g., median age at doctorate, percent with master's degree, percent female) with population profiles derived from the National Research Council's Doctorate Records File and the Office of Education's *Earned Degrees Conferred* shows excellent correspondence. Our *Dissertation Abstracts*-based sample is representative of the 1969–1970 Ph.D. cohort in six fields: sociology, psychology, zoology, biochemistry, physics, and electrical engineering (see Table 1 for sample n per field). The six fields were chosen to span the National Research Council's categories of science-based doctorate holders, although we claim neither that any field is representative of a major category (e.g., physics of the "physical sciences," including mathematics, chemistry, and earth sciences), nor that any set of fields is representative of "science." When respondents were compared with nonrespondents and with those in our original sample not effectively addressed (e.g., foreign addresses, no addresses, undeliverable), no significant differences on such measures as Ph.D. institution prestige or geographical region emerged. Likewise, our initial publication counts showed a nonsignificant difference (t test) between respondents (mean = 6.46, median = 3.62) and nonrespondents (mean = 5.96, median = 2.54), but a significant one with the nonaddressed group (mean = 4.81, median = 2.12). Our respondents thus seem to overrepresent slightly those active in research. We also note that these "objective" publication counts, derived from searching the *Science Citation Index*, greatly underestimate our final tallies augmented by respondent review of our listing and subsequent rechecking as we compiled citation counts (overall sample publication mean = 9.66 vs. original tally mean of 6.46—a distressing shortfall of 33%).

REFERENCES

Astin, Alexander W.
1971 Predicting Academic Performance in College. New York: Free Press.

Barnes, S. B. and R. G. Dolby
1970 "The scientific ethos: a deviant viewpoint." European Journal of Sociology 11:3–25.

Bourdieu, Pierre
1975 "The specificity of the scientific field and the social conditions of the progress of reason." Social Science Information 14:19–47.

Brown, David G.
1967 The Mobile Professors. Washington, D.C.: American Council on Education.

Bucher, R. and J. Stelling
1977 Becoming Professional. Beverly Hills: Sage.

Caplow, T. and R. J. McGee
1958 The Academic Marketplace. New York: Basic Books.

Cartter, A. M.
1966 An Assessment of Quality in Graduate Education. Washington, D.C.: American Council on Education.

Chubin, Daryl E.
1974 "Sociological manpower and womanpower: sex differences in career patterns of two cohorts of American doctorate sociologists." The American Sociologist (May):83–92.
forthcoming "Mentors and students: social heritability in science." In D. E. Chubin, Social Trappings of Knowledge.

Crowley, C. J. and D. E. Chubin
1976 "The occupational structure of science: a log-linear analysis of the inter-sectoral mobility of American sociologists." The Sociological Quarterly 17:197–217.

Hargens, Lowell L.
1969 "Patterns of mobility of new Ph.D.s among American academic institutions." Sociology of Education 42:18–37.
1978 "Theory and method in the sociology of science." Pp. 121–39 in J. Gaston (ed.), Sociology of Science. San Francisco: Jossey-Bass.

Harmon, Lindsey R.
1963 "The development of a criterion of scientific competence." In C. W. Taylor and F. Barron (eds.), Scientific Creativity: Its Recognition and Development. New York: Wiley.

Knorr, Karin D.
forthcoming "Scientific communities or variable transscientific fields? A critique of quasi-economic models of science." Social Studies of Science 11.
Krohn, Roger G.
1971 The Social Shaping of Science: Institutions, Ideology, and Careers in Science. Westport, Conn.: Greenwood.
Long, J. Scott
1978 "Productivity and academic position in the scientific career." American Sociological Review 43:889–908.
Long, J. Scott, Paul D. Allison, and Robert McGinnis
1979 "Entrance into the academic career." American Sociological Review 44:816–30.
Marsh, John F., Jr. and Frank P. Stafford
1967 "The effects of values on pecuniary behavior: the case of academicians." American Sociological Review 32:740–54.
Merton, Robert K.
1942 "Science and technology in a democratic order." Journal of Legal and Political Sociology 1:115–26.
1965 "The ambivalence of scientists." Pp. 112–32 in N. Kaplan (ed.), Science and Society. Chicago: Rand McNally.
Mitroff, Ian I.
1974 "Norms and counter-norms in a select group of the Apollo moon scientists: a case study of the ambivalence of scientists." American Sociological Review 39:579–95.
Mulkay, M. J.
1974 "Methodology in the sociology of science: some reflections on the study of radio astronomy." Social Science Information 13:107–19.
1976 "Norms and ideology in science." Social Science Information 15:637–56.
1980 "Interpretation and the use of rules: the case of the norms of science." Transactions of the New York Academy of Sciences 39:111–25.
Nelkin, Dorothy
1975 "Changing images of science: new pressures on old stereotypes." Newsletter 14 of the Program on Public Conceptions of Science, Harvard University: 21–31.
Porter, Alan L.
1977 "Citation analysis: queries and caveats." Social Studies of Science 7:257–67.
Porter, Alan L. and D. Wolfle
1975 "Utility of the doctoral dissertation." American Psychologist 30:1054–61.
Reskin, Barbara F.
1978 "Scientific productivity, sex, and location in the institution of science." American Journal of Sociology 83:1235–43.
1979 "Academic sponsorship and scientists' careers." Sociology of Education 52:129–46.
Roose, K. D. and C. J. Andersen
1970 A Rating of Graduate Programs. Washington, D.C.: American Council on Education.
Stehr, Nico
1978 "The ethos of science revisited: social and cognitive norms." Pp. 172–96 in J. Gaston (ed.), The Sociology of Science. San Francisco: Jossey-Bass.
Studer, K. E. and D. E. Chubin
1980 The Cancer Mission: Social Contexts of Biomedical Research. Beverly Hills: Sage.
Whitley, Richard
1977 "The sociology of scientific work and the history of scientific developments." Pp. 21–50 in S. S. Blume (ed.), Perspectives in the Sociology of Science. Chichester: Wiley.
Woolgar, S. W.
1976 "Writing an intellectual history of scientific development: the use of discovery accounts." Social Studies of Science 6:395–422.
Zuckerman, Harriet
1977 Scientific Elite: Nobel Laureates in the United States. New York: Free Press.
Zuckerman, Harriet and R. K. Merton
1972 "Age, aging, and age structure of science." Pp. 292–356 in M. W. Riley et al. (eds.), A Theory of Age Stratification. Vol. 3, Aging and Society. New York: Russell Sage.

REPLY TO CHUBIN, PORTER, AND BOECKMANN*

J. Scott Long
Washington State University

Robert McGinnis
Paul D. Allison
Cornell University

Chubin, Porter, and Boeckmann's comment attempts to clarify the limits to which our earlier research can be generalized (Long, 1978; Long et al., 1979) and to augment our study of biochemists with comparative analyses of six fields. In the process they attempt both to demonstrate a lack of replicability of our findings and to suggest future directions of research in the social stratification of science. While we strongly support their suggestion that additional, comparative studies of the scientific career are necessary—indeed essential—we cannot agree with their claims about our position or even that their findings are comparable to ours.

Chubin et al. chastise us for studying academic biochemists and generalizing to all fields of science, over all periods of time, and for all organizational contexts of science. This criticism is unfounded. First, the titles of both papers include the qualifying adjective "academic" to make clear that our intention was not to generalize to nonacademic science. The

* Direct all communications to J. Scott Long, Department of Sociology, Washington State University, Pullman, WA. 99164.

The authors would like to thank Lowell Hargens and Eugene Rosa for their suggestions.

need to distinguish between the academic and nonacademic sectors and among various work activities (which Chubin et al. have not done) is a clear concern of our research program (see Long and McGinnis, 1981a [this issue]). Second, to indicate our generalization from biochemistry to all fields, they quote, "Although these results are for a single scientific field, we see nothing peculiar about biochemistry that would raise doubts about their generalizability (Long et al., 1979:829)." In our article, the quoted sentence is followed by evidence that biochemistry should not be unique and a report that a larger study was in progress. In the absence of contrary evidence, we still do not see any reason why our results cannot be generalized to other physical and biological sciences. However, we strongly encourage *comparable,* longitudinal, cross-disciplinary studies to test this hypothesis. Third, Chubin et al. criticize us for not comparing those of our sample with pre-Sputnik degrees with those with post-Sputnik degrees. We are aware of Sputnik's effects on American science. Indeed, the study was initially designed to include cohorts from before and after Sputnik, but analyses of covariance failed to find any significant differences between these cohorts. We do not claim that certain marginal distributions, such as the percentage of graduates in postdoctoral study or entering academic employment, are constant across time or discipline. Our claim is limited to the processes that affect scientists in similar organizational contexts.

Chubin et al. provide what they believe to be evidence that conflicts with our findings. To demonstrate a contradiction, however, it is necessary that the work be comparable. Unfortunately, because of problems in methodology and errors in specification, the work of Chubin et al. neither supports nor contradicts our results.

If we were to assume that their variables and specifications are comparable to ours, the unstandardized regression coefficients that they present still might not be appropriate. Sociologists generally consider unstandardized coefficients as appropriate for assessing invariance of causal processes across populations. However, an overly rigid acceptance of this practice can be misleading. Indeed, Hargens (1976:251) argues that:

> This acceptance is unwarranted, because there are instances when it seems reasonable to postulate that the allocation of rewards within a social system is based on an individual's abilities or contributions *relative to those of the other members of the social system.* . . . Such a causal process might be invariant across a number of populations, but, given common raw score measures of rewards and performances in the populations studied, the unstandardized coefficients would not reflect this invariance.

The social system of science is the example that Hargens chose to illustrate his argument. While some may disagree with Hargens' argument that standardized coefficients may be structural coefficients, and hence appropriate for comparisons across populations, in comparative analyses of the stratification system in science both standardized and unstandardized coefficients should be considered. Thus, even without the flaws noted below, Chubin et al.'s use of unstandardized coefficients is an inappropriate basis for supporting a claim of lack of replicability of our findings.

Chubin et al. measure key variables in substantively different ways, further attenuating the comparability of results. For our measure of the prestige of the postdoctoral fellowship, they substitute a dummy variable indicating whether the student had a fellowship. We find that those two variables behave quite differently: having a fellowship is not comparable to the prestige of the fellowship. Their measure of the eminence of the mentor is the student's response to the question: "At the time of your dissertation how prominent was your supervisor within his or her area of specialization?" Four response categories were possible, ranging from "Renowned" to "Not prominent." Both the validity of this variable and its treatment as an interval level measure must be questioned. It is clearly not comparable to our measure of the citations received by the mentor. Any differences between their estimates and ours of the effects of these variables clearly can be due to differences in measurement.

Chubin et al. argue that our linkage between mentor stature and first job must be challenged, but they do not even have a measure of the prestige of the first job, and hence their arguments are tenuous. Nonetheless, we are sympathetic with the need to consider further the relationship between the mentor and the student's academic job placement. Indeed, we (Long and McGinnis, 1981b) have undertaken further data collection and analyses assessing the effect of the mentor on the academic career. Variables related to the mentor's eminence, productivity, academic rank, and collaboration with the student have been analyzed, with results that support the findings presented in Long et al. (1979). Our major clarification of the earlier findings is that it is primarily the mentor's productivity, independent of eminence (cf. Reskin, 1979:131), that affects job placement. Further, this effect operates primarily for those students who collaborate with their mentor.

Chubin et al.'s comparisons of regressions on later productivity are based on differently specified models and possibly inappropriate groups of scientists. The major finding of Long (1978) in this regard is that the prestige of the academic job is essential for explaining later productivity. For this reason, analyses of later productivity were restricted to those who did not change institutions, to avoid the confounding effects of departmental changes. Chubin et al. did not restrict their analysis to those academic faculty who do not change departments, and it is not clear whether their sample of academic scientists includes nonfaculty positions. Most importantly, they do not include any measure of the prestige of the employing department. In effect, they ignored the major result of the research they are trying to replicate. The effect of the doctoral department operates indirectly through its influence on initial job placement. Accordingly, it is not surprising that they are not able to "sustain" our finding that Ph.D. prestige does not affect later productivity independently of current departmental prestige. If we exclude prestige of the employing department from our analyses (as they have done), the effect of the doctoral department emerges. Chubin et al. have misspecified our model in such a way that the effects they find are expected, but nonetheless incorrect. Furthermore, cross-disciplinary support for our findings is provided by Hargens and Hagstrom (1981), who include a measure of the prestige of the employing department, for a sample of physicists, chemists, biologists, and mathematicians.

Besides excluding the critical variable from our specification, Chubin et al. included two variables that are not in our model, dropped two additional variables that are in our model, and measured two other variables in substantially different ways. Under these circumstances, comparisons are clearly not appropriate.

In summary, our analyses did not try to generalize to nonacademic science. We still believe that our findings may be applied to other fields in the biological and natural sciences, but welcome and encourage additional, comparative studies with which to test this hypothesis. Replications are clearly necessary in the sociology of science, but such replications must be, in fact, replications. One must be certain that "operationally somewhat different" measures are not simply different measures. When specific processes are compared, key variables must be included. In short, such replications must be comparable if they are to contribute to our understanding. Chubin et al. have not provided this contribution.

REFERENCES

Chubin, D. E., A. L. Porter, and M. E. Boeckmann
1981 "Career patterns of scientists: comment on Long et al., ASR, October 1979." American Sociological Review 46:488–96.

Hargens, Lowell L.
1976 "A note on standardized coefficients as structural parameters." Sociological Methods and Research 5:247–56.

Hargens, Lowell L. and Warren O. Hagstrom
1981 "Consensus and status attainment patterns in scientific disciplines." Working Paper. Indiana University.

Long, J. Scott
1978 "Productivity and academic position in the scientific career." American Sociological Review 43:889–908.

Long, J. Scott, Paul D. Allison, and Robert McGinnis
1979 "Entrance into the academic career." American Sociological Review 44:816–30.

Long, J. Scott and Robert McGinnis
1981a "Organizational context and scientific productivity." American Sociological Review 46:422–42.
1981b "The effects of the mentor on the academic career." Paper presented at the Pacific Sociological Association. March, 1981.

Reskin, Barbara
1979 "Academic sponsorship and scientists' careers." Sociology of Education 52:129–46.

CORRECTION

FIREARMS OWNERSHIP FOR SPORT AND PROTECTION: TWO NOT SO DIVERGENT MODELS*

(CORRECTION TO "FIREARMS OWNERSHIP FOR SPORT AND PROTECTION: TWO DIVERGENT MODELS," LIZOTTE AND BORDUA, *ASR*, APRIL 1980)

ALAN J. LIZOTTE
Indiana University

DAVID J. BORDUA
CAROLYN S. WHITE
University of Illinois, Urbana-Champaign

Introduction

The results of logistic equations reported here replace those in Lizotte and Bordua (1980), which were in error. Equations predicting firearms ownership for sport and for protection were not based on 714 cases as was reported earlier. Rather, they were estimated using only the first 200 cases in the data file. Documentation for the Nerlove-Press modified logistic response model program (LOGLIN31) states that the program can process up to 1200 cases. Indeed, output from the program reported that all cases (714) had been processed. But this is a bit of programming legerdemain. That is, the program reports back the number listed on the number of cases statement, regardless of how many cases were actually processed. Closer inspection of the dimensions of matrices within the program revealed the flaw.

We modified the LOGLIN31 (Cyber version) program to accept 16 independent variables and 714 cases. Since the saturated protection equation required more than 16 independent variables, we screened the initial equation using Probit. Using a liberal .3 level of exclusion for insignificant variables left 13 significant variables. We then proceeded in a step down fashion using LOGLIN31 to estimate the reduced equation with a .05 significance level. We further searched for a preferred model on the basis of chi-squared improvements for equations with variables excluded one at a time.

In the following sections we reestimate the equations predicting gun ownership for sport and for protection. The coding of all variables, the theory employed, and the logic of the analysis remain unchanged. The results do not. OLS equations which were originally reported were not affected by the error, and all footnotes in Lizotte and Bordua (1980) still apply.

Analysis and Results: The Sporting Model

Figure 1 shows the final estimated model for gun ownership for sport. Insignificant variables have been dropped from the model. Unstandardized coefficients are used to measure effects. Sportmags, sex, age at first gun, age,[1] trained, Fhunt, size, and education are the only significant determinants of gun ownership for sport. Sex, age, county hunting, licenses, size 16, and parent's gun ownership also have indirect effects through age at first gun on sporting gun ownership. That is, males, respondents with gun-owning parents, and respondents who live in high–hunting license counties all tend to be young first-time owners. Additionally, respondents who lived in small towns when 16 years old tend to have acquired a gun at an early age. This translates into a high probability of gun ownership for sport in later life.

The corrected equation in Table 3 shows that gun ownership for sport is the result of a sporting gun culture. That is, the model meets our criteria of behavior, socialization, and contact among members, all of which are independent of situation. Significant numbers of people exhibit a behavior which is different from the larger culture—they own guns for sport. In addition, we have located the mechanism of socialization into the subculture. The intergenerational transmittance is carried out via family socialization through parents' gun ownership and age at first gun. There is also

* Direct all correspondence to Alan J. Lizotte, Department of Sociology, Ballantine Hall, Indiana University, Bloomington, Indiana 47405.

We are indebted to Douglas A. Smith for estimating equations for the supporting Probit analysis.

[1] Age at first gun is coded with current age if the respondent never owned a gun. Controlling for age in equations where age at first gun is either endogenous or exogenous purges age at first gun of the age component. However, this procedure does not allow a distinction between the methodological effect of age on gun ownership and the variable's true effect. To test for an age effect on gun ownership we dichotomized at first gun (18 or younger). This method shows a strong effect of age at first gun on sporting ownership. Also, age is not a significant predictor of sporting gun ownership. This suggests that the effect of age reported here is a spurious methodological effect of no substantive meaning.

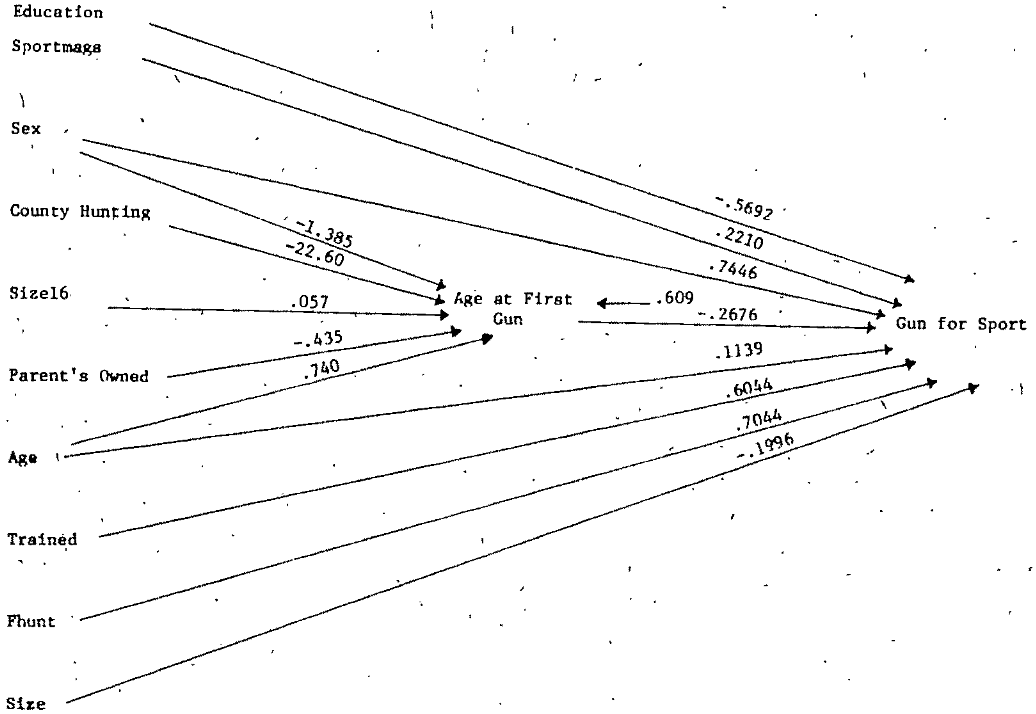


Figure 1. A Path Model of Firearms Ownership for Sport (OLS and Logistic Estimation Procedures Are Used Where Appropriate)

strong evidence of contact among members of the sporting culture. Respondents are likely to own guns when household members hunt together, when the county hunting-rate is high, and when they have been trained to use guns. The contact can be more symbolic. When households maintain contact with the sporting culture via sporting magazine subscriptions, the probability of gun ownership is high. Finally, situation operates independent of this contact and socialization. Individuals are more likely to own guns when the size of the place in which they live is small. Highly educated people tend not to own guns for sport. Additionally, age is positively related to ownership.[2]

Analysis and Results: The Protection Model

We now turn to estimating the corrected equation for firearms ownership for protection. Equations predicting Copcor, victim, perceived crime, fear, and homedef were originally estimated using OLS. Therefore, they remain unchanged. Table 4 and Figure 2 show that fear of crime, age (older), sex (males), race (blacks), size (smaller), friend's gun ownership for protection, and violent experiences[3] all predict respondents' gun ownership for protection. Of course, all of the indirect effects through fear which were originally discussed under "*Other Findings in the Protection Model*" operate on gun ownership for protection.

To this point there is only partial evidence of a subculture of protective gun ownership. We can locate a group of protective gun owners and there is contact among members of the group (friend's protective ownership) which is independent of other situational factors. However, there is no evidence of socialization. Additionally, there is evidence that violent experiences predict protection ownership, but violent attitudes (Vatt1 and Vatt2) do not. The contention that there is a subculture of vio-

[2] Equations with dichotomous endogenous variables were also estimated using probit analysis. The probit equations are very similar to the logistic equations reported here.

[3] When victimization and violent experiences are both included in the protection equation they both become insignificant. If either is excluded the other becomes significant. We report the violent experiences model because it provides a *slightly* better fit than victimization. However, we can not distinguish between the two on statistical criteria. Because of this, caution should be used when interpreting the effect of violent experiences on protective ownership. Upon request we will provide interested readers with the protection equation including victimization and excluding violent experiences.

Tables 3 and 4. Logistic Equations Predicting Gun Ownership for Sport and Gun Ownership for Protection (Standard Errors Are in Parentheses)

Predictor	Table 3 Gun for Sport	Table 4 Gun for Protection
Trained	.6044	
	(.1720)	
Fhunt	.7044	
	(.1914)	
Sportmags	.2210	
	(.09898)	
Age at first gun	−.2676	
	(.0498)	
Education	−.5692	
	(.1862)	
Sex	.7446	.4113
	(.1795)	(.1738)
Age	.1139	.1648
	(.04573)	(.05446)
Size	−.1996	−.1628
	(.06192)	(.07939)
Race		−.6713
		(.2756)
Friend's gun		.6877
		(.1925)
Violent experiences		.2408
		(.08756)
Fear of crime		.4736
		(.1684)
Constant	−.6201	−2.720
	(.3396)	(.4973)
−2 log λ	595.155	704.934
D.F.	9	8

lence centered on protective ownership is not fully supported.

A question remains about the effects of family socialization on protective ownership. Two possibilities exist. First, family socialization variables could have direct effects on protective ownership. Second, family socialization variables could have indirect effects on protective ownership through sporting ownership. If the latter is the case, a subculture of protection could be accepted on grounds that socialization occurs through participation in the subculture of sporting gun use. We can see in Table 5 that the first possibility—direct effects of family socialization on protection ownership—does not obtain. The second possibility does. Ownership for sport has a direct effect on protective ownership and vice versa.[4] Family socialization affects protection ownership indirectly through sport ownership.

Table 5 shows corrected equations for gun ownership for sport with protection ownership used as a predictor, and gun ownership for protection with sport ownership used as a predictor. All variables which were significant in

[4] Causality can not be imputed between sport and protection ownership. That is, we can not determine whether or not (1) sport ownership leads to protective ownership; (2) the reciprocal; or (3) both occur. We leave such questions to more ambitious research. This research only shows that overlap exists, and socialization into protection ownership could obtain through sporting ownership.

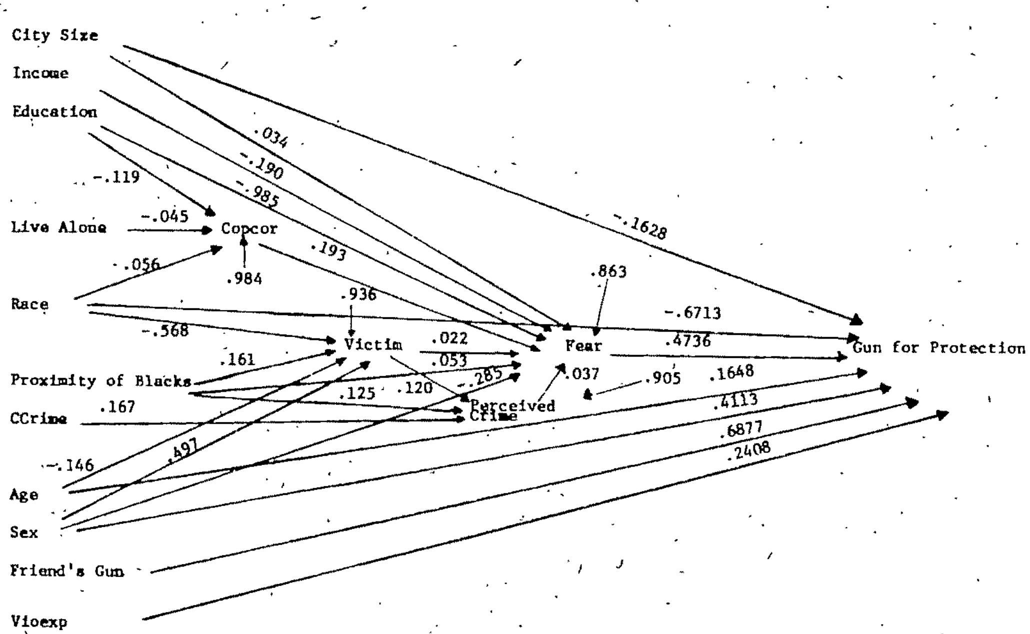


Figure 2. A Path Model of Firearms Ownership for Protection (OLS and Logistic Estimation Procedures Are Used Where Appropriate)

Table 5. Logistic Equations of Joint Probabilities (Conditional Estimators) for Owning a Gun for Sport and Owning gun for Protection (Standard Errors Are in Parentheses)

Predictor	Gun for Sport	Gun for Protection
Constant	−.3309	−1.973*
	(.4683)	(.4915)
Trained	.6170*	.3753
	(.1634)	(.1951)
Fhunt	.6785*	.001504
	(.1892)	(.2279)
Sportmags	.2084*	.1361
	(.1020)	(.1101)
Sex	.7247*	−.1095
	(.1700)	(.1643)
Race	.1398	−.8327*
	(.2470)	(.2609)
Age at first gun	−.2629*	−.005010
	(.04552)	(.06449)
Age	.08446	.1763*
	(.04729)	(.06246)
Education	−.4824*	−.3982*
	(.2024)	(.1890)
Size	−.1843*	−.09834*
	(.06425)	(.04837)
Friend's gun	−.3377*	.8025*
	(.09937)	(.1852)
Violent experiences	.07363	.1414
	(.06618)	(.07331)
Fear of crime	−.01533	.4442*
	(.08791)	(.1524)
Gun for sport	—	.3931*
		(.1083)
Gun for protection	.3781*	—
	(.1023)	
$-2 \log \lambda$	608.603	730.973
D.F.	14	14

* Significant at the .05 level.

either equation individually are included in the conditional estimation equations. For symmetry, the conditioning endogenous variable is coded −1 for no gun, and 1 for gun. For example, the gun for protection equation answers the question: What is the probability of owning a gun for protection given the probability that a gun is owned for sport? The answer is that owning a gun for protection is, in part, a function of owning a gun for sport. Conversely, owning a gun for sport is, in part, a function of owning a gun for protection.

It is important to note that the indirect effect of family socialization through sporting ownership to protection ownership is only one of many effects. Independent of this socialization effect, situation plays a strong role in the ownership of guns for protection. That is, crime, perceived crime, victimization, and fear of crime all have either direct or indirect effects on protective ownership, and these effects are independent of socialization and the protective subculture. Unfortunately, we do not know the specific reason for a respondent's owning a protection gun. For example, farmers may own guns for protection from wild dogs and other predatory animals, while those living in fear of crime probably own for protection from humans. This would explain the findings on protective ownership by such divergent groups as rural respondents and those respondents who have been victimized and are fearful of crime. Although we have no way of knowing, significant overlap between protective and sporting ownership might not exist if we were able to exclude those respondents who own guns for protection from animals. It is also important to note that in the joint estimation equation (Table 5), violent experiences no longer predict protection ownership.

Summary

There is no doubt that socialization plays a strong role in the sporting use of guns. Parents' gun ownership is a strong predictor of age at first gun, and age at first gun is a strong predictor of owning a sporting gun. This indicates socialization. Further, family hunting and subscriptions to magazines both predict sporting-gun ownership. This shows evidence of contact among members. These findings support the notion of a subculture of gun use.

Not all of the subcultural indicators were significant predictors of firearms ownership for sport. Several studies have found a strong relation between veteran status and firearms ownership (see Newton and Zimring, 1968). This relation was proven spurious. Veterans are no more likely than others to own guns because of their military training.

Gun ownership for protection also behaves, in part, subculturally. We found evidence of socialization into the protection subculture, since some respondents who own guns for protection are socialized via the sporting culture. However, we have no way of knowing whether protection ownership involves only protection from humans or partly also from animals. There is also contact among protection owners. That is, respondents are likely to own guns for protection when they have friends who own guns for protection.

Since violent attitudes did not predict gun ownership for protection and violent experiences did not predict in the joint estimation equation (Table 5), we conclude that a subculture of violence centered on protective ownership does not exist. We do not mean to imply that this study is conclusive evidence that subcultures of violence do not exist. Other

measures of violent attitudes and behavior may be better predictors.

Trust in the police and the courts, victimization, and perceived crime all have indirect effects on protective ownership. These indirect effects are through fear of crime. Protective gun ownership was not found to be the logical extension of a home-defense orientation. Interestingly, relative to men, women are more likely to own guns for protection than for sport. Racist attitudes do not predict owning a gun for protection.

REFERENCES

Lizotte, Alan J. and David J. bordua
1980 "Firearms Ownership for Sport and Protection: two divergent models." American Sociological Review 45:229–44.

Nerlove, Marc and S. James Press
1973 Univariate and Multivariate Log-Linear and Logistic Models. Santa Monica: Rand.

Newton, George D. and Franklin E. Zimring
1968 Firearms and Violence in American Life. Washington, D.C.: U.S. Government Printing Office.

American Sociological Review

Volume 46 October 1981 Number 5

SCOPE AND MISSION OF AMERICAN SOCIOLOGICAL REVIEW: The *American Sociological Review* publishes work of interest to the discipline in general, new theoretical developments, results of research that advance our understanding of the most fundamental social processes, and important methodological innovations. Like other publications of the Association, emphasis is given to exceptional quality. Unlike the more specialized journals of the Association, the *American Sociological Review's* primary focus is on whatever has the most general bearing on the knowledge of society.

Concerning manuscripts, address: Editor, *American Sociological Review*, Department of Sociology, University of Illinois, Urbana, Illinois 61801. (217) 333-6515.

The *American Sociological Review* (ISSN 0003–1224) is published at 49 Sheridan Avenue Albany, New York, bimonthly in February, April, June, August, October, and December.

Applications for permission to quote from this journal should be addressed to the Executive Office

Annual membership dues of the American Sociological Association, \$20–\$120 (\$10 of this amount may cover a subscription to the *American Sociological Review*); Associate, \$25; International Associate, \$25; Student Associate, \$12.

Application for membership and payment of dues should be made to the Executive Office. Subscription rate for members, \$10; nonmembers, \$18; institutions and libraries, \$36; nonmember students, \$12. Single issue, \$4.

New subscriptions and renewals will be entered on a calendar year basis only.

Change of address: Six weeks advance notice to the Executive Office and old address as well as new are necessary for change of subscriber's address.

Concerning advertising, changes of address, and subscriptions, address: Executive Office, American Sociological Association, 1722 N Street, N.W., Washington, D.C. 20036.

Claims for undelivered copies must be made within the month following the regular month of publication. The publishers will supply the missing copies when losses have been sustained in transit and when the reserve stock will permit.

Second class postage paid at Washington, D.C. and additional mailing offices.

(Articles in the REVIEW are indexed in the Social Sciences Index, Psychological Abstracts, Sociological Abstracts, Ayer's Guide, University Microfilms, Abstracts for Social Workers, International Political Science Abstracts, and United States Political Science Documents.)

AMERICAN SOCIOLOGICAL REVIEW

Official Journal of the American Sociological Association

OCTOBER, 1981 VOLUME 46, NUMBER 5

(Continued on next page)

COMMENTS AND REPLIES

Notice to Contributors

A processing fee of $10 is required for each paper submitted, except for comments; such fees will be waived for student members of ASA. A check or money order, made payable to the American Sociological Association, should accompany each submission. The fee must be paid in order for the review process to begin.

The length of an *ASR* paper is generally ten (10) printed pages or thirty (30) typed pages. However, the *ASR* is happy to accept manuscripts up to fifty (50) typed pages in length for review. We also are interested in receiving manuscripts for a Research and Notes Section.

To permit anonymity in the review of manuscripts, keep identifying material out of the manuscript. Attach a cover page giving authorship, institutional affiliation and acknowledgments, and provide only the title as identification on the manuscript and abstract.

ASA Multiple Submissions Policy: Submission of a manuscript to a professional journal clearly implies commitment to publish in that journal. The competition for journal space requires a great deal of time and effort on the part of editorial readers whose only compensation for this service is the opportunity to read papers prior to publication and the gratification associated with discharge of professional obligation. For these reasons, *the American Sociological Association regards submission of a manuscript to a professional journal while that paper is under review by another journal as unacceptable.*

SUBMISSION AND PREPARATION OF MANUSCRIPTS

1. Submit four (4) copies and retain the original for your files. Copies may be Xerox, mimeograph or multilith, but not carbons.
2. Enclose a stamped self-addressed postcard for acknowledgment of manuscript receipt. *Manuscripts will not be returned.* Please do not send your original copy.
3. All copy must be typed, doublespaced (including indented material, footnotes and references) on 8½ by 11 inch white opaque paper. Lines must not exceed six (6) inches. Margins must be a minimum of one inch.
4. Include four (4) copies of an abstract of no more than 150 words.
5. Type each table on a separate page. Insert a location note at the appropriate place in the text, e.g., "Table 2 about here." Send copies, retain originals.
6. Figures must be drawn in India ink on white paper. Send copies, retain originals.
7. Clarify all symbols with notes in the margins of the manuscript. Circle these and all other explanatory notes not intended for printing.
8. Footnotes should not be placed in the body of the text. Type them (doublespaced) and attach them as a separate appendix to the text. Number them consecutively throughout the text. Footnotes are to be kept to a minimum and used only for substantive observations. Source citations are made within the text rather than in footnotes.
9. Acknowledgments, credits and grant numbers are placed on the title page with an asterisk.

REFERENCE FORMAT

A. *In the text:* All source references are to be identified at the appropriate point in the text by the last name of the author, year of publication and pagination where needed. Identify subsequent citations of the same source in the same way as the first, not using "*ibid.*," "*op. cit.*," or "*loc. cit.*" Examples follow:

1. If author's name is in the text, follow it with year in parentheses. [". . . Duncan (1959) . . ."]
2. If author's name is not in the text, insert, in parentheses, the last name and year, separated by a comma. [". . . (cf. Gouldner, 1963) . . ."]
3. Pagination (without "p." or "pp.") follows year of publication after a colon. [". . . Kuhn (1970:71)."]
4. Give both last names for dual authors; for more than two use "et al." in the text. When two authors have the same last name, use identifying initials in the text. For institutional authorship, supply minimum identification from the beginning of the complete citation. [". . . (U.S. Bureau of the Census, 1963:117) . . ."]
5. Separate a series of references with semicolons and enclose them within a single pair of parentheses. [". . . (Burgess, 1968; Marwell et al., 1971; Cohen, 1962) . . ."]

B. *In the appendix:* List all source citations by author, and within author by year of publication, in an appendix titled "References." The reference appendix must be complete and include all references in the text. The use of "et al." is not acceptable in the appendix; list the names of all authors. (See A. 4. for text format.)
If there is more than one reference to the same author and year, distinguish them by the letters, a, b, etc. added to the year. [". . . Levy (1965a:331) . . ."] Give the publisher's name in as brief a form as is fully intelligible. For example, John A. Wiley and Sons should be "Wiley." If the cited material is unpublished, use "forthcoming" with name of journal or publisher; otherwise use "unpublished."
Use no underlining, italics or abbreviations.
Examples follow:

1. *Books:*	Jud, Gerald J., Edgar W. Mills, Jr. and Genevieve Walters Burch	
	1970	Ex-Pastors. Philadelphia: Pilgrim Press.
	U.S. Bureau of the Census	
	1960	Characteristics of Population. Volume 1. Washington, D.C.: U.S. Government Printing Office.
	Bernard, Claude	
	[1865] 1957	An Introduction to the Study of Experimental Medicine. Tr. Henry Copley Greene, New York: Dover.
2. *Periodicals:*	Conger, Rand	
	Forthcoming	"The effects of positive feedback on direction and amount of verbalization in a social setting." Pacific Sociological Review.
	Merton, Robert K.	
	1963a	"The ambivalence of scientists." Bulletin of The Johns Hopkins Hospital 112:77–97.
	1963b	"Resistance to the systematic study of multiple discoveries in science." European Journal of Sociology 4:237–82.
3. *Collections:*	Davie, M.	
	1938	"The pattern of urban growth." Pp. 133–61 in G. Murdock (ed.), Studies in the Science of Society. New Haven: Yale University Press.

See recent issues for further examples.

Revised 1980 ASR

CONTRIBUTORS TO THIS ISSUE

■ LAWRENCE E. COHEN (Social Inequality and Predatory Criminal Victimization) is Associate Professor, Department of Sociology, University of Texas at Austin. He is studying the relationship between unemployment and crime and the forecasting of crime.

■ JAMES R. KLUEGEL is Associate Professor, Department of Sociology, University of Illinois at Urbana-Champaign. Besides studying the social determinants of victimization risk, he is engaged in research on the structure, determinants, and consequences of stratification ideology, based on a survey of Americans' beliefs about social inequality.

■ KENNETH C. LAND is Professor, Department of Sociology, and Research Associate, Population Research Center, at University of Texas at Austin. He is constructing annual demographic accounts for the United States from 1940 to 1980, building time-series models of monthly birth and death rates, and studying the relationship between fluctuations in unemployment and crime rates. He is coeditor, with F. Thomas Jester, of *Social Accounting Systems: Essays on the State of the Art* (Academic Press, forthcoming).

■ MICHAEL D. WIATROWSKI (Social Control Theory and Delinquency) is Assistant Professor, Department of Criminology, Florida Atlantic University. His current interests include the social ecology of homicide and the relationships between educational processes and delinquency.

■ DAVID B. GRISWOLD, Assistant Professor, Department of Sociology at Florida Atlantic University is studying sentencing policy.

■ MARY K. ROBERTS is Research Associate, University of Florida. Her main interest is vocational psychology.

■ DENISE C. GOTTFREDSON (Black-White Differences in the Educational Attainment Process) is Associate Research Scientist, Center for Social Organization of Schools, The Johns Hopkins University. She is involved in a national evaluation of a major initiative to prevent delinquency through school-based intervention programs. She is also studying the determinants of school disruption and of individual-level delinquency and socioeconomic attainment.

■ A. WADE SMITH (Racial Tolerance as a Function of Group Position) is Assistant Professor, Department of Sociology, Arizona State University, and Research Associate at the Center for the Study of American Pluralism, National Opinion Research Center. He is studying the efficiency with which different sampling techniques locate additional black respondents to the 1982 General Social Survey and is developing time-series models of the formation, organization, and dissolution of multi-adult families.

■ GILLIAN STEVENS (Social Mobility and Fertility), a Ph.D. candidate at the University of Wisconsin, Madison, will soon be joining the faculty at the University of Illinois at Urbana-Champaign. Her interests include minority language maintenance in the United States and minority language-group patterns of endogamy and exogamy.

■ GLENN R. CARROLL (Dynamics of Organizational Expansion in National Systems of Education) is Assistant Professor, Brown University, and Guest Assistant Professor, Zentrum fur Umfragen, Methoden und Analysen in Mannheim. His current research focuses on the organizational ecology of newspapers.

■ GLENN V. FUGUITT (Community Structure in Response to Population Growth and Decline) is Professor, Department of Rural Sociology, University of Wisconsin–Madison. He studies population and business changes in the American village and migration and population redistribution in nonmetropolitan areas.

■ JOHN D. KASARDA is Professor and Chairman, Department of Sociology, University of North Carolina at Chapel Hill. He is involved with research on national urban policies and with the study of ecological models of organizational structure.

■ MYRON BOOR (Effects of United States Presidential Elections on Suicide and Other Causes of Death) is Assistant Professor, Section of Psychiatry and Human Behavior, Brown University, and Research Psychologist, Rhode Island Hospital. His interests include thanatology and the demography of suicide.

■ KARL L. ALEXANDER (Measure for Measure) is Associate Professor, Department of Social Relations, The Johns Hopkins University. He is engaged in research on race, gender, and SES differences in college graduation prospects and, with Martha Cook, on leisure activities as sources of vocational preparation. He remains interested in refinements of social-psychological models of the status-attainment process.

■ AARON M. PALLAS is a Ph.D. candidate,

Department of Social Relations, The Johns Hopkins University. He is studying, with Karl Alexander, the impact of high-school coursework on standardized test performance and sex differences in mathematics achievement.

■ MARTHA A. COOK is Assistant Professor, Department of Sociology, Case Western Reserve University. She is investigating curriculum placement as an organizational decision made within the comprehensive high school, conceptualizing the school as a service organization with students filling a complex client role.

■ HART M. NELSON (Religious Conformity in an Age of Disbelief) is Professor and Chair of Sociology and Head of Rural Sociology, Louisiana State University. He is studying the religious socialization of children and youth and the impacts of seminaries upon the attitudes of Presbyterian clergy.

■ DOUGLAS S. MASSEY (Social Class and Ethnic Segregation) is Assistant Professor, Population Studies Center and Department of Sociology, University of Pennsylvania. He is engaged in analysis of patterns of residential succession in Hispanic neighborhoods and a study of the relationship between education and U.S.-bound migration from a Mexican town.

■ KENNETH A. BOLLEN (The Dimension(s) of Democracy) is Senior Research Scientist, Societal Analysis Department, General Motors Research Laboratories. He is studying imitative suicides that follow publicized suicide stories and is involved with several cross-national projects on the political and demographic consequences of economic development.

■ BURKE D. GRANDJEAN is Assistant Professor, Department of Sociology, University of Virginia. He is studying stratification in the United States civil service.

■ CHARLES M. JUDD (Political Involvement and Attitude Structure in the General Public) is Associate Professor in the Department of Psychology, University of Colorado at Boulder. Research interests include attitude structure and measurement and methods of applied social research. He is author, with D. A. Kenny, of *Estimating the Effect of Social Interventions* (Cambridge University Press, 1981).

■ JON A. KROSNICK is a Ph.D. candidate, Department of Psychology, University of Michigan. His research interests include the causes of adolescent substance abuse, attribution theory in social psychology, and attitude structure.

■ MICHAEL A. MILBURN is Assistant Professor, Psychology Department, University of Massachusetts–Boston. He is studying the structure of attitude/belief systems and the communication content of soap operas.

ERRATA

The following errors appeared in Jackson, and Carroll, "Race and the War on Crime," *ASR*, June 1981.

On page 295, right-hand column, sixth line from the bottom, the parenthetical phrase should read

(population in 1970)

On page 302, in the note to Figure 2, two plus signs should be changed to minus signs:

NOTE: Capital Expenditures $\hat{Y} = -.26961X + .01072X^2 - .00011X^3$.
Expenditures for Salaries and Operations $\hat{Y} = -.34040X + .02784X^2 - .00035X^3$.

SOCIAL INEQUALITY AND PREDATORY CRIMINAL VICTIMIZATION: AN EXPOSITION AND TEST OF A FORMAL THEORY*

LAWRENCE E. COHEN
University of Texas at Austin
JAMES R. KLUEGEL
University of Illinois at Urbana-Champaign
KENNETH C. LAND
University of Texas at Austin

This study systematically tests a formal theory of how certain dimensions of social stratification—income, race, and age—relate to risk of predatory criminal victimization. An opportunity theory of criminal victimization is proposed, focusing on the mediating role played by five risk factors: exposure, guardianship, proximity to potential offenders, attractiveness of potential targets, and definitional properties of specific crimes themselves. Propositions are derived pertaining to the bivariate and multivariate-partial (main) effects expected from the theory and tested in analyses based on a representative sample of the U.S. population for the crimes of assault, burglary, and personal larceny.

These data indicate that the relationship between the dimensions of social stratification and the offenses studied here is complex, and that, other things being equal, those usually thought to be most vulnerable economically and socially—the poor, the nonwhite, the old—are not *the most likely victims of crime. Race has little direct effect on victimization risk, while age is inversely related to each type of crime at both the bivariate and multivariate levels of analyses. The findings are largely consistent with the proposed theory.*

INTRODUCTION

The relationships of social inequality to criminal behavior and to the arrest, prosecution, and sentencing of alleged criminal offenders are among the most frequently studied topics in American criminology (for excellent reviews of several of these studies, see Nettler, 1978, 1979). Many citizens and criminologists alike believe that the disadvantaged are less adequately insulated than the advantaged from conditions that stimulate crimes and that the disadvantaged receive less favorable treatment from the criminal justice system. Similar beliefs exist with respect to protection from predatory criminals. Quinney (1975:129), for example, contends that

> Even for conventional crimes, the victims are those who are already oppressed in the society. Except for auto theft, the victims of all the major conventional crimes are disproportionately in the lower income levels . . . except for larceny, the highest rates of victimization are for blacks. The lower class and blacks, in particular, are major victims of conventional crimes.

If this statement is valid, it cannot be explained by better protection afforded the wealthy by the differential concentration of law enforcement resources alone, since high crime–low income areas are generally the most heavily patrolled within police jurisdictions.

Freedom from criminal victimization is a social good much like any other social good. As research on the fear of crime

---

* Direct all correspondence to Lawrence E. Cohen, Department of Sociology, University of Texas, Austin, Texas 78712.

The research reported here was supported, in part, by National Science Foundation grant SES 77-13261 and by National Institute of Mental Health grant 1-R01MH3117-02. The data used in this paper were made available by the Inter-University Consortium for Political and Social Research. We thank Michael Gottfredson for helpful comments and suggestions and David Cantor for assistance in the preparation and analysis of the data. We, of course, are solely responsible for the contents of this paper.

Table 1. Proportions of Persons or Households (for Burglary) Victimized for Crimes of Assault, Burglary, and Personal Larceny by Categories of Income, Race, and Age

	Assault	Burglary	Personal Larceny
Income			
1. <7,500	.013 (67,585)	.041 (39,308)	.031 (67,585)
2. 7,500–24,999	.009 (125,034)	.036 (55,553)	.041 (125,034)
3. 25,000+	.008 (16,910)	.041 (6,715)	.056 (16,910)
TOTAL	.010 (209,529)	.038 (101,576)	.039 (209,529)
Race			
1. White	.010 (187,382)	.036 (90,677)	.040 (187,382)
2. Black	.011 (22,147)	.054 (10,899)	.035 (22,147)
Age			
1. 16–29	.019 (68,506)	.058 (19,971)	.062 (68,506)
2. 30–59	.007 (98,298)	.037 (53,582)	.034 (98,298)
3. 60+	.002 (42,725)	.025 (28,023)	.014 (42,725)

SOURCE: National Crime Surveys for 1974 and 1977, merged.
NOTE: The difference between whites and blacks in proportion victimized for the crime of assault is not statistically significant ($p>.10$). All other differences are statistically significant ($p<.001$).
The category n is given in parentheses.

(see, for example, Garofalo, 1979) and on public opinions about crime (Hindelang, 1975) attests, the possibility of being victimized deeply troubles many people. Until quite recently, however, criminologists lacked any substantial understanding of the nature and distribution of various forms of criminal victimization. Due to continuing large scale studies of victimization (see, for example, U.S. Department of Justice, 1974a, 1974b, 1974c, 1976, 1977), we are now beginning to acquire the resources necessary to map out the distribution of victimization and to explore the factors that increase or decrease exposure to risk for different types of crime.

In this paper, we analyze the relationship between certain aspects of social inequality and the risk of criminal victimization. We begin by examining the simple bivariate distributions of victimization risk, for several of the most serious and frequently occurring crimes, by three important dimensions of social stratification: income, race, and age.[1] Next, we present a formal theory of predatory victimization risk. This theory offers an empirically verifiable explanation for the observed bivarate and multivariate-partial (main effect) relationships between these dimensions of social inequality and risk of criminal victimization. Using data obtained from a representative sample of the U.S. population, we then test, through log-linear analysis, the postulates and theorems of this formal theory.

DIMENSIONS OF STRATIFICATION AND CRIMINAL VICTIMIZATION

As a starting point, Table 1 presents bivariate distributions of victimization for assault, burglary, and personal larceny by income, race, and age, using data from large, representative national surveys of the U.S. (detailed information on these surveys will be provided below).[2] The

[1] Consistent with much recent work on the aged in American society (see, for example, Riley et al., 1972), we view age as an important dimension of stratification—that is, as a factor, like race and income, that systematically affects the distribution of social goods among persons.

[2] Certain serious crimes (such as robbery and rape) were not included in this analysis because their frequency of occurrence does not permit as extensive multivariate analyses as are required for the test of our formal theory. According to the definitions used in the NCS, an assault is defined as an unlawful physical attack, whether aggravated or simple, excluding such attacks accompanied by theft or attempted theft. A burglary is defined as forcibly entering a residence, or attempting to forcibly enter, or unlawfully entering such a premise without force to commit a felony. Personal larceny offenses include all (nonauto) thefts or attempted thefts from a person (rather than a commercial establishment) in which no force or threat of force is used, occurring: at or in a vacation home, hotel, or motel; inside a commercial structure such as a store, restaurant, bank, gas station, public conveyance, or station; inside an office, factory, or warehouse; on the street, in a park, field, playground, school ground, or parking lot; or inside a school.

gross differences in victimization risk for the different crimes among the different income, race, and age categories are social facts that a theory of predatory victimization should explain.

It is apparent from Table 1 that a simple unidimensional characterization of how inequality relates to victimization is not possible. It is not the case that those usually thought to be economically and socially most vulnerable—the poor, the nonwhite, the old—are consistently the most likely victims of crime. For example, the risk of victimization does not occur along a strict economic dimension. While for personal larceny the risk of victimization increases monotonically with income, the risk of burglary victimization is slightly higher among both the poorest and the most affluent citizens than it is for those in the middle income group. In contrast, assault victimization is inversely related to income. Furthermore, while nonwhites have a higher overall risk of burglary, there is virtually no difference by race in assault victimization risk, and, for the crime of personal larceny, it is whites who appear to have the higher risk of victimization. Finally, contrary to expectations on the basis of a strict vulnerability argument, risk is inversely related to age for all types of victimization.

A FORMAL THEORY OF PREDATORY VICTIMIZATION

The explanation of criminal victimization risk we offer is rooted in a recently emerging theoretical perspective which we extend and formalize (for compatible theoretical statements, see Land and Felson, 1976; Hindelang, Gottfredson, and Garofalo, 1978: Chap. 11; Cohen and Felson, 1979; Cohen, Felson and Land, 1980; Cohen and Cantor, 1980). Broadly, this perspective, which may be referred to as *the opportunity model of predatory victimization*, considers the time-space relationships in which victimization risk is greatest. The risk of criminal victimization is seen as largely dependent on the lifestyle and routine activities of persons that bring them and/or their property into direct contact with potential offenders in the absence of capable guardians who could potentially prevent the occurrence of a crime.

In our judgment, the key to understanding why income, race, and age appear to affect the likelihood of victimization in the ways they do is to focus on the mediating role played by five factors: exposure, guardianship, proximity to potential offenders, attractiveness of potential targets, and definitional properties of specific crimes themselves. The first two factors (exposure and guardianship) are dimensions of what Hindelang and his colleagues (1978: Ch. 11) generally call *lifestyle*. They offer an explanation of how dimensions of inequality affect the risk of victimization by estimating the effects of these dimensions on lifestyle and the effects of variations in the latter on victimization risk. While Hindelang, Gottfredson, and Garofalo (1978) have contributed substantially by laying a groundwork for a general theory of victimization, we believe that their framework overemphasizes the role of lifestyles in mediating the effects of social inequality on victimization risk and that the propositions they offer lack the specificity necessary for building a formal theory.

Definitions

Our formal theory linking dimensions of social inequality to criminal victimization commences with definitions of five factors which are strongly related to risk:

Exposure: the physical visibility and accessibility of persons or objects to potential offenders at any given time or place.

Proximity: the physical distance between areas where potential targets of crime reside and areas where relatively large populations of potential offenders are found.[3]

[3] The factor of proximity differs from that of exposure in that the former is a physical relational property pertaining to physical distances between residential locations of populations of potential targets and potential offenders, whereas the latter pertains to variations in physical visibility and accessibility of potential targets (persons or objects) to potential offenders as determined by personal characteristics of the potential targets. Our theory postulates, for example, that of two potential targets whose personal characteristics (e.g., household structure, employment status) imply equal exposure, but who dif-

Guardianship: the effectiveness of persons (e.g., housewives, neighbors, pedestrians, private security guards, law enforcement officers) or objects (e.g., burglar alarms, locks, barred windows) in preventing violations from occurring, either by their presence alone or by some sort of direct or indirect action.

Target Attractiveness: the material or symbolic desirability of persons or property targets to potential offenders, as well as the perceived inertia of a target against illegal treatment (i.e., the weight, size, and attached or locked features of property inhibiting its illegal removal and the physical capacity of persons to resist attack). We further differentiate target attractiveness on the basis of whether the motivation to take action against a specific person or object is primarily *instrumental* (i.e., the act is a means of acquiring something one desires) or *expressive* (i.e., the act of attacking a person or stealing property is the only reward sought in doing so).

Definitional Properties of Specific Crimes: the features of specific crimes that act to constrain strictly instrumental actions by potential offenders. For example, many larcenies are less difficult to commit and require less knowledge of victim routine activities than do burglaries (for an explanation of this fact, see the related discussion of Assumption 5 in the next section). Such constraints limit the ability of potential burglary offenders to consistently act against targets that would maximize their economic gain, thus requiring them to seek out less attractive targets.

Assumptions

We specify five assumptions about links between the risk factors defined above and the likelihood of criminal victimization.

(1) *Exposure.* All else equal, an increase in exposure leads to an increase in victimization risk.

In order for a predatory criminal victimization to occur, a motivated offender must come into contact with a potential victim and/or the victim's property. The more frequently such contact occurs, the more opportunity there will be for an offender to act against a potential victim and/or his/her property.

(2) *Guardianship.* All else equal, offenders prefer targets that are less well-guarded to those that are more well-guarded. Therefore, the greater the guardianship, the less the risk of criminal victimization.

Elsewhere, Cohen, Felson, and Land (1980:98) refer to this assumption as the "rational behavior" postulate.

(3) *Proximity.* All else equal, the closer the residential proximity of potential targets to relatively large populations of motivated offenders, the greater the risk of criminal victimization.

This assumption is based on what has been called the "gravity law of distance and social interaction" (see Olsson, 1964). As in the case of Assumption 1, Assumption 3 is based on the definitional property of predatory crime that contact between a potential offender and a potential victim is required; our analyses are confined to those predatory violations involving direct physical contact between at least one offender and at least one person or object, which that offender attempts to steal or damage. Being in spatial proximity to motivated offenders tends to increase the frequency of regular contact of offenders with potential victims and thus increases the opportunity for offenders to act against a potential victim and/or his/her property.

(4) *Attractiveness.* All else equal, if a crime is motivated by instrumental ends, the greater the attractiveness of a target, the greater the risk of victimization.

This assumption requires that, in the aggregate, property crime is substantially the result of instrumental motives. We do not deny that property crime may also be motivated by expressive ends (thrill-seeking, peer-group acceptance, etc.); all that is needed for this assumption to be applicable, however, is for one to grant that a significant percentage of property crime is committed primarily for economic gain.

(5) *Properties of Crimes.* The strength of the partial effects of exposure, guardianship, and proximity on victimization risk depends

---

fer in proximity, the person in closer proximity to populations of potential offenders will have a higher risk of victimization for reasons indicated under Assumption 3 in the text.

upon the degree to which properties of crimes themselves constrain strictly instrumental action. Specifically, the more constrained strictly instrumental action is, the stronger will be the effects of exposure, guardianship, and proximity on victimization risk relative to the effect of target attractiveness.

This assumption does not apply to the crime of assault since it seems largely motivated by expressive ends (Hindelang et al., 1978:265). On the other hand, there are two reasons why personal larceny is a crime that involves fewer constraints on strictly instrumental action than does burglary (thus personal larceny victimization risk is more strongly affected by target attractiveness than is burglary victimization risk). First, personal larceny—theft outside the household—is more open to strict instrumental behavior on the part of offenders because the value of property being stolen is generally more visible to the offender than is the case for burglary. Burglary, more often than personal larceny, constitutes a calculated gamble taken by an offender based on a guess about the potential value of property contained within a residence. Borrowing a useful cliche, in many cases "appearances will be deceiving," thus giving a significant proportion of burglary victimizations a seemingly noninstrumental quality. Second, the successful completion of a burglary requires, on the average, more information about a potential victim's routine activities than does a personal larceny (see Walsh, 1974:404; Reppetto, 1974:17). At a minimum, information about whether or not a domicile is unoccupied is desirable. For these reasons, exposure and guardianship play especially important roles in determining the risk of burglary victimization. Exposure is important since in the interaction with potential offenders, requisite information about potential victims' routine activities and the value of their property will be obtained. Proximity is also relevant in these regards. The more a potential victim's daily activities permit ready scrutiny by potential offenders, the more easily they obtain reliable information about guardianship (occupancy patterns) and the value of property. The closer a person lives to potential offenders, the more easily one's daily activities can be scrutinized.

The series of graphs in Figure 1 summarize the relationships postulated in these five assumptions. These graphs will be used as an aid for the derivation of theorems in a subsequent section of this paper. A "+" or "–" is used to indicate the postulated direction of each relationship.[4] The strengths of the relationships are represented by the slopes of the lines—the steeper the slope, the stronger the postulated relationship.

Links between Inequality and Risks

As a means of organizing what prior research tells us about the links between dimensions of inequality and our stated risk factors, we present three summary principles. In other theories, these principles could be the objects of explanation; here, they form a convenient point of departure, characterizing existing social structure.

Principle of Homogamy: To the degree that persons share sociodemographic characteristics with potential offenders they are more likely to interact socially with such potential offenders, thus increasing the risk factor of exposure.

This principle has been stressed in the analyses given by Hindelang et al. (1978:257–9). The match of sociodemographic characteristics is important, because the resulting lifestyle similarity is likely to bring potential offenders and potential victims into direct contact more often than when such characteristics are not shared. Also, the similarity of characteristics provides potential offenders with information about a potential victim's likely activity patterns by virtue of the offender's knowledge of his/her own activity patterns, and by previous direct or indirect interaction among offenders and

[4] Because our postulates are stated in terms of first-order relationships (first-order partial derivatives), the graphs in Figure 1 are linear. This should not be taken to imply that we are unaware of the possibility that some of these relationships may be nonlinear, at least in certain parts of their ranges. Rather, we merely focus our present analysis, which already is long and complicated, on the first-order effects.

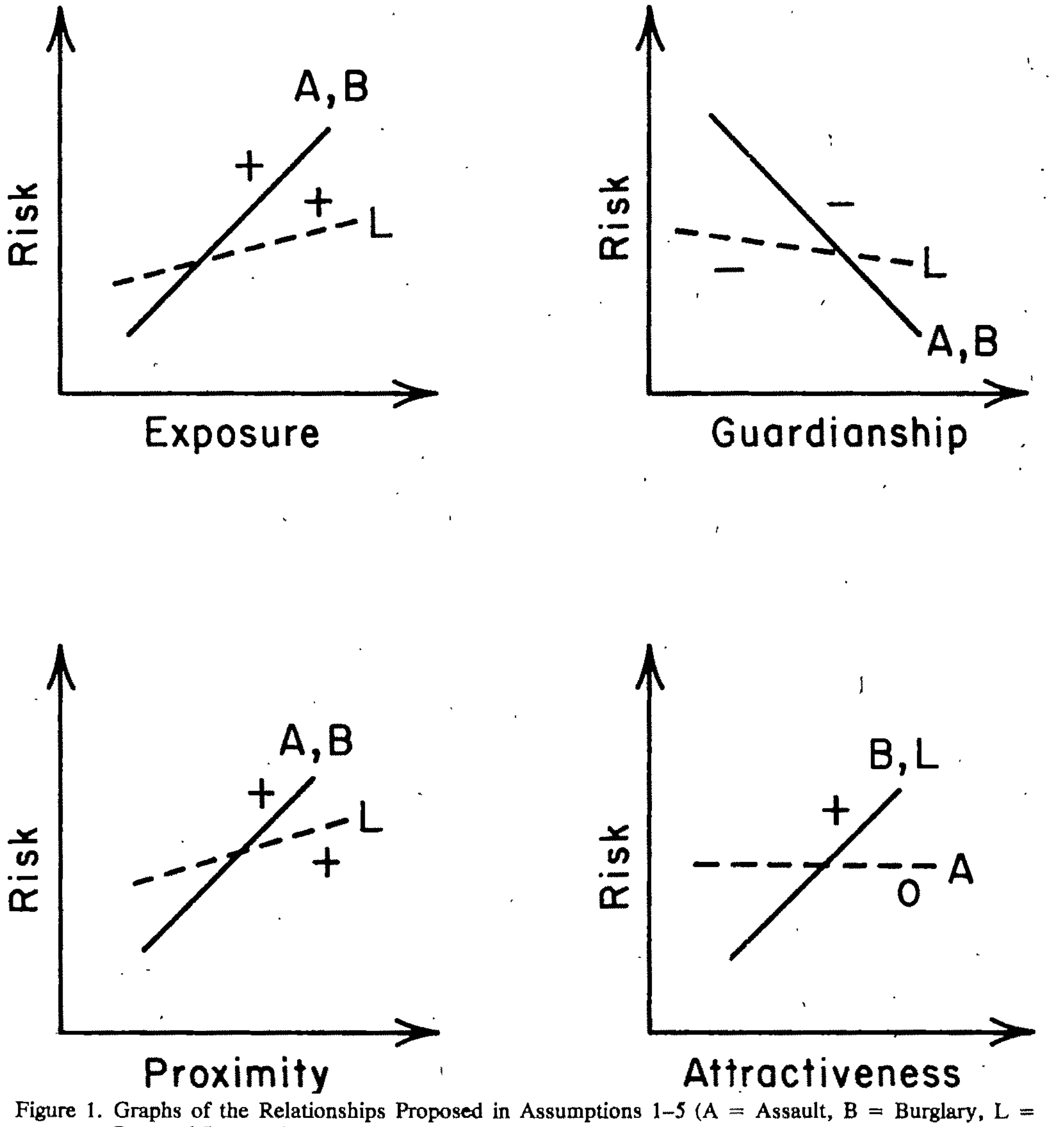


Figure 1. Graphs of the Relationships Proposed in Assumptions 1–5 (A = Assault, B = Burglary, L = Personal Larceny)

their potential victims. Thus, for example, a potential offender is more likely to have "inside knowledge" about a potential victim who shares her/his characteristics—knowing about such things as times when the potential victim is likely to be away from home, features of his/her dwelling, specific property available for theft, and so on.

This principle has clear implications for the relationship of age, race, and income to victimization risk. Prior research shows that Index Crime offenders are disproportionately poor, nonwhite, and young (see Hindelang, 1978). Thus, by the principle of homogamy, the poor, nonwhites, and the young have higher exposure to potential offenders than do other income, race, and age categories.

Principle of Dependence of Guardianship on Inequality Dimensions: To the degree that a person's lifestyle places her/him and related others (through primary group ties) in routine contact with a potential target (property or a person), that person and the related others will provide stronger guardianship for the target (Cohen, Felson, and Land, 1980).

This principle implies that the degree to which a person is routinely in contact with her/his property or significant others will,

on the average, serve to increase deterrence capabilities against predatory criminals. If more than one related persons are in such routine contact, even stronger guardianship is provided, since a potential offender may be able to neutralize the retaliatory response of a single individual but doing so for two or more related persons is much more problematic. We emphasize "persons related by primary group ties" as such persons will be more likely to have a mutual interest in each other's welfare (Cooley, 1909).

Individuals' lifestyles differ in both their amount of contact with their potential property targets and the number of persons routinely in contact with their property. For example, being married and residing with one's spouse by itself increases guardianship, because two related persons will be routinely in contact with potential property targets. Being married with one person routinely at home (e.g., a housewife) provides even stronger guardianship, since there will be both two related persons routinely in contact with property and one person who will be in greater daily contact with property at home than would be the case if both the husband and wife worked outside of the home. Prior research suggests that income, race, and age may be systematically related to lifestyle differences that affect the strength of guardianship (Hindelang et al., 1978:259–62) in the following ways: an increase in income leads to stronger guardianship patterns, being nonwhite is associated with weaker guardianship patterns, and growing older is associated with stronger guardianship patterns.

We expect income to be related to guardianship patterns because high incomes are associated with the means to purchase relatively effective security devices or services available to better guard one's person and property. In addition, higher incomes, all else equal, make it more economically feasible to marry, and also increase one's attractiveness as a marriage partner—with marriage increasing the number of persons routinely attached to potential targets, thereby providing for improved guardianship capabilities (Cohen and Felson, 1979; Cohen, Felson, and Land, 1980). Racial differences in marriage and family structure also have implications for guardianship. Since nonwhite households more often consist of a currently unmarried adult (never married, divorced, or separated) than white households (U.S. Bureau of Census, 1979: Tables 1 and 2), their risk of victimization due to lesser guardianship capability will, in general, be higher than for whites. Furthermore, race, by virtue of its association with income, is also related to guardianship for reasons specified above. Finally, age relates to guardianship in that younger persons (adolescents and young adults) spend considerably more time outside the home than do older persons. DeGrazia (1961:121–4) for example, found that activities centered away from home decline steadily after the early twenties to almost zero for persons over sixty, and Chapin (1974:113) reports that the very young and the elderly have more "free time" than middle-aged persons and that the elderly disproportionately spend their free time watching television and resting. By their mid-twenties, people are more likely to be married and adopt home-centered activities (U.S. Bureau of Census, 1973), which leads to a positive correlation of age and guardianship.

Principle of Residential Segregation by Inequality Dimensions: People tend to live in areas that are homogeneous with regard to race, income, and age.

Here one needs to do no more than point to the large body of accumulated literature on residential segregation (see, for example, Farley, 1977; Taeuber and Taeuber, 1965). Given the disproportionate representation of Index Crime offenders among the poor, young, and nonwhite noted earlier, implications for the relationship of income, race, and age to the risk factor of proximity clearly follow. The lower the income, the closer the proximity. The lower the age, the closer the proximity. Nonwhites are in closer proximity than whites.

Before concluding this section, a few comments about the relationship of income, race, and age to target attractiveness are needed. We assume that, all else equal, a person's income provides a good

direct operational indicator of his/her attractiveness as a potential property target. However, we know of no theoretical or factual basis on which to argue that race and age have net relationships to target attractiveness (although it is tempting to speculate that older persons, because they are generally least able to resist or retaliate, make especially attractive targets for such expressive crimes as assault). Hence, all else being equal, we assume null relationships of race and age to target attractiveness.

Figure 2 presents a series of graphs (in the same form as those of Figure 1) that summarize the foregoing principles concerning the effects of income, race, and age on risk factors in predatory criminal victimization.

Propositions

On the basis of the definitions, assumptions, principles, and associated discussions presented above and summarized in Figures 1 and 2, the following propositions

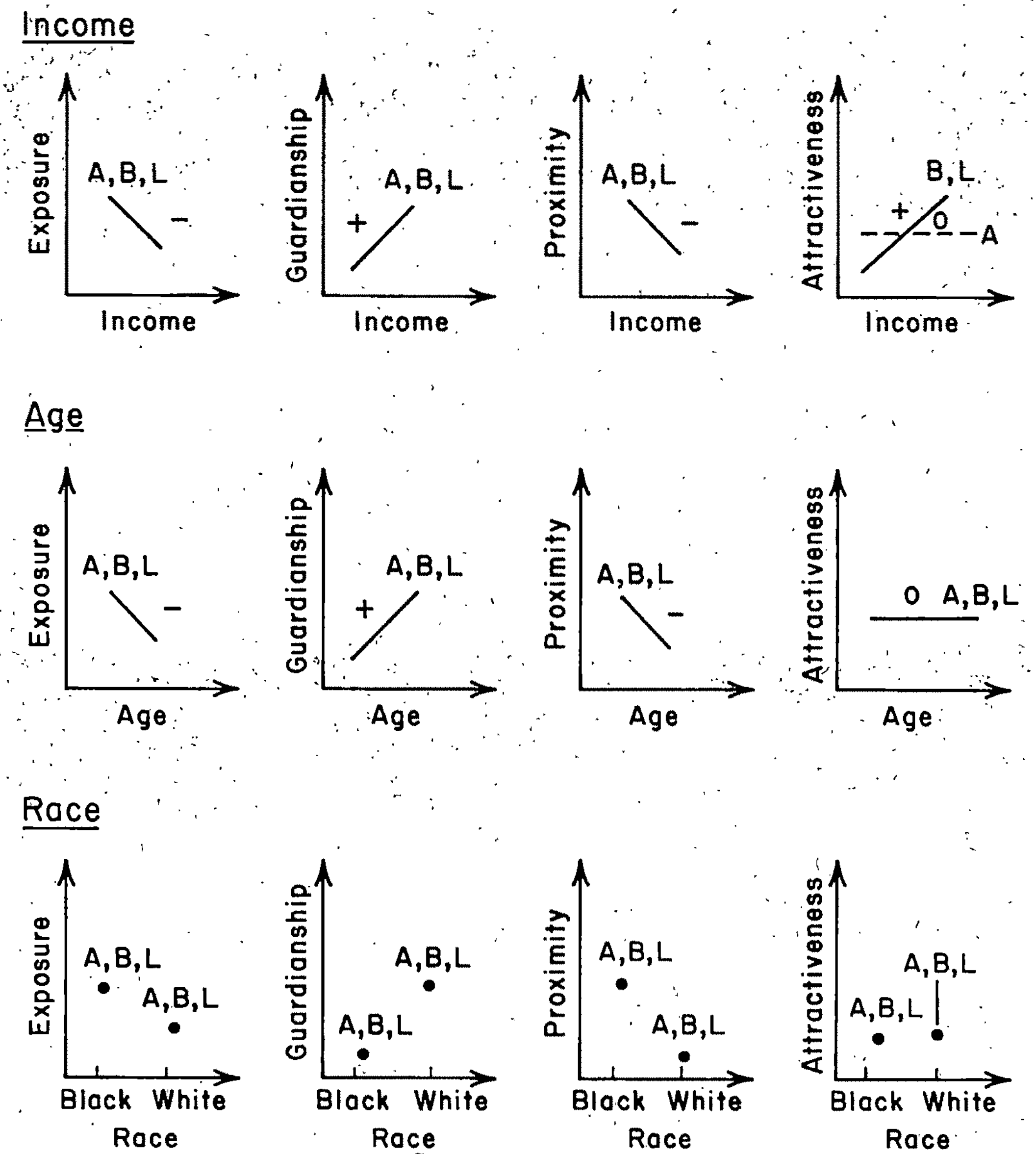


Figure 2. Graphs of Relationships Between Age, Income, and Race and the Four Risk Factors (B = Burglary, A = Assault, L = Personal Larceny)

pertaining to the bivariate and partial distributions by income, race, and age of the types of victimization considered herein can be derived.

(1) The bivariate distribution of assault victimization by income is negative—the higher the income, the lower the risk of assault victimization. This proposition is derived by combining the partial effects of income on the victimization risk factors (exposure, guardianship, proximity, attractiveness) for assault exhibited in Figure 2 with the postulated relationships of the risk factors to the risk of assault shown in Figure 1. Since the reduced-form effects of income on risk of assault victimization through the first three risk factors (exposure, guardianship, proximity) are negative and that through the last risk factor (attractiveness) is null, we obtain the stated total or marginal relationship of income to the risk of assault victimization. (All of the following propositions on marginal effects may be similarly derived by combining the relationships graphed in Figures 1 and 2, but we shall not do so in such laborious detail.)

(2) Income will have two opposing partial effects on burglary victimization risk. On the one hand, increases in income will lead to lower risk of victimization through exposure, guardianship, and proximity. On the other hand, increases in income will increase risk of victimization due to attractiveness. The net result will tend to be either no bivariate differences by income categories or a curvilinear relationship such that increases in income lead to decreases in victimization risk up to a point of high income, where risk will increase with increasing income.

(3) Income will have a weak negative effect on personal larceny risk through guardianship, proximity, and exposure, such that the higher the income, the lower the victimization risk. But income will have a strong positive effect on personal larceny risk through attractiveness. On balance, then, the bivariate distribution of personal larceny victimization risk will be positive—i.e., increases in income will be associated with increases in personal larceny victimization risk.

(4) For assault, burglary, and personal larceny, the bivariate distributions by age will be negative. Increasing age will be associated with decreasing victimization risk for each of the three types of crime studied here.

(5) Nonwhites will have a higher bivariate risk of assault and burglary victimization than will whites.

(6) Nonwhites will have a slightly lower bivariate personal larceny victimization risk than will whites. There will be two effects here. On the basis of the effects of race through exposure, guardianship, and proximity, nonwhites will tend to have higher victimization risk than whites. However, through the association of race with income (and its strong effects on the risk of personal larceny victimization), nonwhites will tend to have a lower risk. The latter effect will overtake the former because of the stronger relationship of income to personal larceny victimization risk than those of exposure, guardianship, and proximity.

(7) Holding lifestyle and proximity constant, income, race, and age will not have direct partial effects on the risk of assault victimization.

(8) Holding lifestyle and proximity constant, income will have a direct partial effect on the risk of burglary and personal larceny victimization.

(9) Holding lifestyle, proximity, and attractiveness (income) constant, race and age will not have direct partial effects on the risk of burglary and personal larceny victimization.

Propositions 1 through 6 refer to the *bivariate distributions* of victimization risk for assault, burglary, and personal larceny by income, race, and age respectively. A reexamination of observed bivariate relationships (see Table 1) between the three dimensions of inequality and the three types of crimes studied here shows that our theory accounts quite well for these relationships.

Propositions 7 through 9 refer to the *direct partial effects* of income, race, and age, controlling for the associations among these variables themselves and for their impacts through the risk factors of lifestyle (exposure and guardianship) and proximity. In addition to testing these propositions, an evaluation of our theory ideally should also include direct tests of

the assumptions. In the remainder of this paper, we present the results of such a multivariate analysis of the effects of income, race, and age and measures of risk factors on the risk of criminal victimization for the crimes of assault, burglary, and personal larceny.

STUDY DESIGN

Sample Characteristics

The analyses described below are based on data from the LEAA National Crime Surveys (NCS) of households within the U.S. (see Penick and Owens, 1976, for a description of this survey). Since each type of predatory criminal victimization is a relatively rare event (see Table 1), their probability density functions are highly skewed. Hence, studies which use types of predatory criminal victimization as dependent variables, such as this one, require a large sample size in order to obtain statistically reliable results. To achieve a sample large enough to be reliable, we employ data from two years of this survey, 1974 and 1977. We have analyzed each sample separately and in merged form. The analyses presented here are based on the merged sample for the two years. With the exception of certain interaction effects discussed later, effects for the separate samples were the same as for the merged sample.

Since, by definition, burglary is a crime committed against a household, while assault and personal larceny are crimes against persons, the unit of analysis differs by type of victimization—household for burglary, and individual for assault and personal larceny. Also, because the overall unit of analysis for these surveys is the household (and all individuals within the household over the age of twelve were interviewed), there is a clustering factor that must be considered before making inferences about true population rates of personal larceny and assault victimization. Since our principal purpose in these analyses is to estimate structural effects and not to present precise population-rate estimates, we will not adjust here for this clustering effect. For the offense of burglary the sample consists of 101,576 households, and for the offenses of personal larceny and assault, the sample consists of 209,529 persons. Cases with missing data were excluded. The households sampled in each year of this survey are divided into six subsamples of approximately 10,000 each. Occupants of each household, twelve years of age or older, are interviewed at six-month intervals for a total of three years. Every six months a new rotation group enters the sample and the eldest existing rotation group is excluded from analysis. Hence, employing the samples from 1974 and 1977 ensures that no household studied here will appear more than once in our analysis.

Variables

Criminal victimization will be defined here as a criminal incident (a burglary, assault, and/or personal larceny) involving one or more offenders and one or more victims in one or more locations within the specified time period, as recorded in the NCS interview. *Income* represents the total family income of the household and is classified into three categories: (1) $7,500 or less; (2) $7,501 to $24,999; and (3) $25,000 or greater.[5] *Race* will be employed here as a dichotomous variable consisting of (1) white, and (2) nonwhite categories.[6] *Age* is classified into three categories: (1) 16–29; (2) 30–59; and (3) 60 and older.[7]

Exposure and guardianship are indexed

---

[5] Analyses were performed using four categories of age and four categories of income, dividing the middle categories reported here for each variable into two smaller categories. No differences among persons in these two middle categories were found, so the simpler three category divisions were used in the final analyses.

[6] Nonwhites consist of blacks and persons of other races who are not white, such as Asians. Persons of Spanish heritage are counted as whites. Persons of racially mixed parentage are classified in the father's racial category. The race of the respondent is determined by observation for members of the same family and by inquiry for households with persons unrelated to the household head.

[7] While data are available for individuals between the ages of twelve and fifteen, we exclude such persons from this study. There is reason to suspect differences in the meaning of reported victimizations (e.g., assault victimization) for this group from that for persons sixteen and over.

here by a proxy measure that we will refer to subsequently as *lifestyle*. Using information obtained from the survey interviews, we have constructed a seven-category variable based on information about respondents' household composition and labor-force statuses. These categories generally designate major differentials in the likelihood of exposure to victimization and in the amount of guardianship provided (see Hindelang et al., 1978:245–64). These categories are (1) not married and employed (students are included among the employed throughout); (2) not married and unemployed; (3) not married and not in the labor force (i.e., unable to work or retired); (4) married, with husband and wife employed and no children; (5) married, with both husband and wife employed with children; (6) married, with head of household employed and wife (or husband) of head not in the labor force; and (7) married, with head of household unemployed. Prior research (Cohen and Felson, 1979; Cohen and Cantor, 1980, 1981) has used marital status and labor-force status as proxies for exposure and guardianship, but, in general, has not employed as extensive an accounting for the total household configuration as we employ here. Two important dimensions underlie this set of categories.

The first is the degree to which one's routine activities are likely to be home-centered. Married people have been found to be more likely on the average to have home-centered activities than those currently not married (see Cohen and Felson, 1979). We expect that those currently married will have a lower risk of criminal victimization due to the lesser exposure and greater guardianship that result from home-centered routine activities. There also may be differences, on the average, within the two broad categories of married and not married in the amount of home-centered activity. Thus, we distinguish between households where both the husband and wife are working and there are no children and other households consisting of married persons. The former households are likely to engage in more activities outside of the home both because of their work status and because of the absence of children (see Cohen and Felson, 1979). In general, extending the logic of the above distinction, we expect that categories 4, 5, and 6 will stand in this order regarding the magnitude of risk and lesser guardianship due to greater exposure to victimization (see also Hindelang et al., 1978, Ch. 11 for a theoretical justification of these empirical facts).

The second important underlying dimension of interest here involves routine activities related to labor-force status. The distinction between being employed and being unemployed is relevant by virtue of the principle of homogamy. Since potential offenders are generally thought to be disproportionately unemployed (Brenner, 1976), the principle of homogamy implies that the unemployed are also more likely to be victims (Hindelang et al., 1978:262–4). As implied above, the labor-force status of all members of the household additionally is relevant since it affects the amount of guardianship of the household that is routinely provided. Finally, we distinguish between being unemployed and being out of the labor force for other reasons (retired or physically unable to work). The category of not married and not in the labor force was created under the assumption that persons in this category would have greater exposure and lesser guardianship than those currently married, but lesser exposure and greater guardianship than the other two categories of those not currently married (see Hindelang et al., 1978:253–62).

Proximity is measured here by a seven-category variable based on the median income of neighborhoods and on features of the geographic area of residence. Respondents were first classified into three groups: residents of central cities or urban areas, residents of noncentral and middle-sized urban areas (population greater than 10,000 and less than 50,000), and residents of small towns and rural areas. Persons within the first two categories were further grouped by the median family income of the neighborhood in which they lived, based on 1970 census data for tracts and enumeration districts (a neighborhood is a group of approximately 4,000 people living in an area about the size of a census tract). Three categories of median neighborhood

income were used: median income less than $7,000, median income greater than $7,000 but less than $20,000, median income of $20,000 or more. On the basis of this classification, the seven-category variable of proximity is as follows: (1) low income, living in central city; (2) middle income, living in central city; (3) high income, living in central city; (4) low income, living in noncentral city; (5) middle income, living in noncentral city; (6) high income, living in noncentral city; (7) living in small town or rural area.

This set of categories reflects both gross distinctions between residential areas commonly thought to be important in victimization risk and, through the use of median neighborhood family income, finer differences among residential areas within these gross categories (see, for example, Hindelang et al., 1978:246–66).

Method of Analysis

To evaluate our explanation of how income, race, and age affect the risk of victimization for the three different crimes studied here, we use log-linear modelling techniques for the analysis of multi-way contingency tables. Although our principal interest is testing the merits of our theory, we employ a broader or exploratory approach to the analysis of these data, rather than a narrower or strictly confirmatory approach. Specifically, we use Goodman's (1972, 1973, 1976) procedures for finding the best-fitting model of how the variables of income, race, age, lifestyle, and proximity combine to influence the risk of criminal victimization for assault, burglary, and personal larceny. This exploratory approach allows us to test the merits of our theory; and it may also reveal facts (e.g., higher-order interactions) for which our theory does not fully account, identifying directions in which future theory construction and research may proceed.

Using Goodman's procedure for selecting the model that most parsimoniously accounts for the associations studied, we fit a series of hypothesized models for the associations in three six-way tables of income by age, by race, by lifestyle, by residential area, and by each crime studied here. Among these, we seek the least complicated model that best accounts for the association between the possible determinants and risk of criminal victimization. Based upon the best-fitting model found in the first stage of our analysis, we present effect coefficients (the log odds) that provide a description of precisely how each hypothesized determinant affects each type of victimization risk. As explicated by Goodman (1976), log-odds coefficients provide analogous information for the case of a dichotomous dependent variable to that provided by regression coefficients in the case of a continuous dependent variable. Finally, the result of model fitting and the effect coefficients are used to further examine the merits of our formal theory.

RESULTS

Model Fitting

Table 2 presents the results of model-fitting tests for each type of victimization risk. The same model-fitting strategy is used for each type of crime. Models 1 through 6 in each of the three panels of Table 2 give tests of models that decrease in complexity from Model 1 (the saturated model) to Model 6 (the independence model). Phrased in the usual analysis of variance terms, Models 1–4 concern interaction effects of decreasing complexity (from 1 to 4), Model 5 (two-way interactions) pertains to the main effects, and Model 6 posits no effects of the independent variables on risk of criminal victimization. Starting with Model 1, each subsequent model deletes one order of interactions, ending with a model with no interactions (Model 6). The likelihood-ratio chi-squared values (denoted by G^2) for tests of the fit of these models (specified under the "Tests of Models" heading in Table 2) are used to start the search for the association between the possible determinants of criminal victimization risk and actual victimization. By calculating the difference between the G^2 value for each model and the G^2 for the model immediately preceding it and the difference in degrees of freedom between these two models, a test of the statistical significance of all interactions of a specific

Table 2. Results of Model Fitting Tests for the Effects of Age(A), Race(R), Income(I), Lifestyle(L), and Proximity(P) On Assault, Burglary, and Personal Larceny Victimization(V)

TESTS OF MODELS			TESTS OF SPECIFIC TERMS			
Model	D.f.	G^2	Terms	G^2 Diff.	D.f.	p
Assault						
1. Saturated model	0	0	The 6-way interaction	31.49	144	>.10
2. All 5-way interactions only	144	31.49	All 5-way interactions	148.63	336	>.10
3. All 4-way interactions only	480	180.12	All 4-way interactions	243.15	280	>.10
4. All 3-way interactions only	760	423.27	All 3-way interactions	222.60	104	<.001
5. All 2-way interactions only	864	645.87	All 2-way interactions	1,452.70	17	<.001
6. Independence model (no effects)	881	2,098.57				
7. (ARILP) (ILV) (LAV) (IAV) (IRV) (IPV) (ARV)	814	478.63				
8. (ARILP) (LAV) (IAV) (ARV) (PV)	846	578.52				
Burglary						
1. Saturated model	0	0	The 6-way interaction	38.73	144	>.10
2. All 5-way interactions only	144	38.73	All 5-way interactions	191.50	336	>.10
3. All 4-way interactions only	480	230.23	All 4-way interactions	227.39	280	>.10
4. All 3-way interactions only	760	457.62	All 3-way interactions	161.17	104	<.001
5. All 2-way interactions only	864	618.79	All 2-way interactions	762.15	17	<.001
6. Independence model (no effects)	881	1,380.84				
7. (ARILP) (LAV) (IAV) (PV) (RV)	848	543.87				
Personal Larceny						
1. Saturated model	0	0	The 6-way interaction	47.66	144	>.10
2. All 5-way interactions only	144	47.66	All 5-way interactions	214.81	336	>.10
3. All 4-way interactions only	480	262.47	All 4-way interactions	274.92	280	>.10
4. All 3-way interactions only	760	537.39	All 3-way interactions	227.59	104	<.001
5. All 2-way interactions only	864	764.98	All 2-way interactions	2,605.00	17	<.001
6. Independence model (no effects)	881	3,369.98				
7. (ARILP) (LAV) (IAV) (IPV) (ARV)	834	638.45				
8. (ARILP) (LAV) (IAV) (ARV)	346	668.57				

order is provided. For example, the difference in G^2 between Model 2 and Model 1 and its associated difference in degrees of freedom provides a test of the statistical significance of the six-way interaction. These significance tests for groups of interactions are given in Table 2 under the heading "Tests of Specific Terms."

Table 2 indicates that, for each of the three types of crime, there are no statistically significant four-, five-, or six-way interactions. Also, for each of the three types of crime, there are significant three-way interactions. Thus, the most parsimonious model for each type of crime contains two-way and three-way effects, but no higher-order interactions.

To complete the specification of the best-fitting model, further tests were run to determine which among the ten possible three-way interactions involving victimization risk were statistically significant. To simplify presentation we omit intermediate results.[8] Model 7 in each panel of Table 2 gives the best-fitting model that includes three-way interactions selected strictly on the basis of statistical significance. In the cases of personal larceny and assault victimizations, additional considerations led us to further simplify Model 7, resulting in the choice of Model 8 for these offenses as the best-fitting and most-reliable model for the determinants of criminal victimization risk. Specifically, for personal larceny and assault victimization, some interactions included in Model 7 were deleted because of small frequencies in certain cells for combined categories of the relevant independent variables. For example, the proximity by income cross-classification contains a cell for the combination of the high family income and the low-income non-

[8] These results are available upon request from the first author.

Table 3. Log-Odds Coefficients for the Two-Way Partial Effects (Main Effects) of Risk Factors on the Three Types of Victimization

	Assault	Burglary	Personal Larceny
Income			
1. <7,500	−.042	−.130	−.370
2. 7,500–24,999	−.164	−.130	−.042
3. 25,000+	.206	.260	.412
Race			
1. White	−.072	−.110	.008
2. Nonwhite	.072	.110	−.008
Age			
1. 16–29	.498	.296	.390
2. 30–59	−.254	−.216	−.116
3. 60+	−.244	−.080	−.274
Lifestyle			
1. Not married, employed	.018	.042	.206
2. Not married, unemployed	1.378	.576	.704
3. Not married, not in labor force	−.156	−.044	−.340
4. Married, both working, no children	−.618	−.394	−.262
5. Married, both working, with children	−.226	.000	−.090
6. Married, head of household working	−.750	−.484	−.448
7. Married, head of household unemployed	.356	.304	.228
Proximity			
1. Central city, low income	.074	.218	.016
2. Central city, medium income	.050	.098	.024
3. Central city, high income	.114	.004	.050
4. Noncentral city, low income	.174	.190	.124
5. Noncentral city, medium income	−.068	−.174	−.020
6. Noncentral city, high income	−.032	−.110	.014
7. Small town and rural areas	−.312	−.226	−.208

central or middle-sized urban area categories that has a frequency of only 14 cases. This strongly suggests that the three-way interactions involving proximity and income may be an artifact of this small cell frequency. A general consideration of the need to be conservative in one's claims for the existence of effects led us to choose Model 8 as the best-fitting and the most conservative representation of how these five risk factors affect risk of personal larceny and assault victimization.[9]

[9] The decision to delete certain statistically significant three-way interactions because of substantive questionability was based on two broad considerations. First, we examined the frequencies in cells for the configuration of categories of the independent variables alone. If, as in the cases of the income-proximity combination and the income-lifestyle combination, very small frequencies were found, we concluded that the relevant three-way interaction is an artifact of this small frequency and the extreme skewness of the dependent variable. Second, we examined the consistency of the three-way interactions both across the two survey years and across the three types of criminal victimization. The three-way interactions reported in the final models exhibited such consistency, while the ones deleted did not.

Effect Coefficients

We have noted that the bivariate relationships displayed in Table 1 offer empirical support for Propositions 1–6. The log-odds coefficients, estimated under the final models, for the partial or multivariate effects are presented in Tables 3 and 4. The two-way interactions are arrayed in Table 3, while the three-way interactions are presented in Table 4. A minus sign (−) indicates that the relative characteristic has less than average odds of being victimized, while a positive value indicates greater than average risk. A coefficient of .000 denotes average risk of victimization.

From these tables, we can make several general points about how income, race, age, lifestyle, and proximity combine to affect the risk of burglary, assault, and personal larceny victimization. We focus initially on the coefficients for the two way interactions and then examine and interpret the three-way interactions. First, it is apparent that, as predicted from Propositions 7 and 9, race has little direct effect on the risk of predatory criminal victimization for each of the three crimes

Table 4. Log-Odds Coefficients for the Three-Way Partial Effects (Interactions) of Risk Factors on the Three Types of Victimization[a]

Lifestyle × Age × Assault

Age	Lifestyle 1	2	3	4	5	6	7
1	.320	−.592	.296	.150	−.164	.264	−.274
2	.344	−.082	.130	.138	−.210	.184	−.504
3	−.664	.674	−.424	−.288	.374	−.450	.778

Lifestyle × Age × Burglary

Age	Lifestyle 1	2	3	4	5	6	7
1	.004	.084	.290	−.016	−.320	.018	−.060
2	.152	−.128	.156	.036	−.136	.186	−.268
3	−.156	.044	−.448	−.020	.456	−.204	.328

Lifestyle × Age × Personal Larceny

Age	Lifestyle 1	2	3	4	5	6	7
1	.154	−.288	.188	.064	−.016	.172	−.272
2	.024	.082	.024	−.074	−.026	.078	−.110
3	−.178	.208	−.212	.010	.042	−.250	.382

Income × Age × Assault

Age	Income 1	2	3
1	.208	−.048	−.160
2	.128	.080	−.206
3	−.336	−.032	.366

Income × Age × Burglary

Age	Income 1	2	3
1	.128	−.050	−.078
2	.046	.026	−.072
3	−.176	.026	.150

Income × Age × Personal Larceny

Age	Income 1	2	3
1	.174	−.076	−.100
2	−.008	.024	−.016
3	−.166	.050	.116

Race × Age × Assault

Age	Race 1	2
1	.252	−.252
2	.120	−.120
3	−.370	.370

Race × Age × Personal Larceny

Age	Race 1	2
1	.150	−.150
2	.070	−.070
3	−.218	.218

[a] Effects are for the odds of being victimized.

(see Table 3). For personal larceny, the main effect of race is virtually zero, and for the other two types of crime it is quite small. Second, the effect of proximity on victimization risk (although stronger than that of race) is generally smaller than each of the effects of income, age, and lifestyle. The clearest contrast among the categories of the proximity variable is between the category of small town and rural areas and all other categories of urban areas. Specifically, controlling for all other variables, persons living in rural areas have a lower overall risk of criminal victimization for each of the three types of crime studied here than do urban dwellers. Third, age and lifestyle have the consistently strongest effects on victimization risk for all three types of crime. For age, the clearest contrast is between persons of ages 29 or younger and those persons 30 years of age or older, with the

former having higher odds of being criminally victimized than the latter. Three subinfluences on victimization risk can be seen by examining the pattern of log-odds coefficients for the lifestyle variables:

> All else equal, persons who are currently married run a lower risk of being victimized for each of the three crimes than those currently unmarried.
>
> Persons who are unemployed, all else equal, run a higher risk of being victimized for each of the three types of crime than do those who are employed or out of the labor force.
>
> Among households where both husband and wife work, having a child seems to increase victimization risk for each type of crime relative to those in this category without children. However, individuals in both types of households, in general, have less than average odds of victimization risk.

Finally, while income has a direct effect on all three types of crime, the log-odds coefficients show this effect to be stronger for personal larceny victimization than for either burglary or assault.

The three-way interactions (Table 4) show that, for all three types of crime, the effect of age is specified (i.e., conditioned) by categories of lifestyle and income. For the crimes of personal larceny and assault, race also serves as a specifying (i.e., conditioning) variable for the effects of age. First, the three-way interaction of lifestyle by age by victimization shows consistently that age has weaker effects on victimization risk among the unemployed and in families where both the husband and wife are employed and have children. However, age has stronger effects on risk of victimization among those unmarried and employed and those unmarried and out of the labor force. Second, the income by age by criminal victimization interaction shows that the effect of age differs by income category. Among the poor, this effect is stronger than that observed among the affluent for each type of victimization. Finally, we note that, for personal larceny and assault (but not burglary) victimization, the effect of age is specified (i.e., conditioned) by race. Age has stronger effects—the risk of being victimized decreases faster with increasing age—among whites than among nonwhites.

DISCUSSION

The multivariate findings provide envidence directly relevant to the validity of the assumptions of our theory. Consistent with each of Assumptions 1–3, for instance, we found that exposure, guardianship, and proximity all have significant partial effects on the risk of predatory victimization.[10] Furthermore, the effect coefficients for the variables of proximity and lifestyle are ordered in terms of magnitude and sign in accord with the expectation predicted by our assumptions. Although the effect coefficients for the categories of proximity are small, they do indicate that proximity to the central city and low-income areas increases predatory victimization risk. In the case of lifestyle (consistent with our previous discussion of guardianship), married persons do face a lesser risk of victimization than do unmarried persons. Also, consistent with our discussion of exposure, unemployed persons run a higher risk of victimization than do employed individuals. The only significant departure from our theoretical expectations is found in the category of persons currently married where both partners work and have children. However, one plausible explanation of this apparent departure from our expectation entirely consistent with our general theory is that such persons have a higher risk of predatory victimization than do married and employed persons without children as a consequence of increased exposure resulting from the association between one's children and their peers. That is, through such associations, potential offenders (friends or acquaintances of the couples' children) may acquire knowledge of attractive targets and routine guardianship patterns which increase the likelihood of victimization among this subgroup of the population.

[10] Strictly speaking, since we have found significant three-way interactions for the hierarchical models estimated from these data, tests for the significance of certain two-way interactions are not possible. However, if we, for the moment, ignore the significant three-way interactions and test for the independent main effects of each of these variables (i.e., base our tests on model 5 each type of victimization), we find that they are statistically significant.

The data also generally support Assumption 4. We assume here that property crime is largely the result of instrumental motives, and Table 3 shows that income is directly related to victimization risk for both personal larceny and burglary. Interestingly, Table 3 also indicates that, all things equal, those persons in the highest income category also have the greatest risk of assault victimization (an offense generally regarded as motivated by expressive concerns). Since the crimes which are defined as assaults in our study exclude attacks accompanied by theft or attempted theft, this finding was not predicted from our theory. A closer inspection of our data, however, indicates that income has a strong effect on risk of assault victimization only among citizens who are in the age category "sixty or older." Relative to the other age by income categories in our study, there are fewer citizens (a total of 1,828) sixty or over with family incomes exceeding twenty-five thousand dollars per year. Therefore, the finding that assault victimization risk is highest for those in the highest income category may not be statistically reliable. Further research with larger samples will be necessary in order to substantiate this relationship.

Finally, we have direct confirmation of our fifth assumption. An examination of the log-odds coefficients in Table 3 show that, as postulated, income has stronger effects on risk of personal larceny victimization than on burglary victimization. That is, the difference in the partial log-odds coefficients between the high and low income categories for personal larceny is roughly twice as large as that for burglary.

On the basis of our definitions, assumptions, principles, and associated discussions, nine propositions were derived pertaining to the bivariate and partial distributions expected from our theory. The data strongly support the expected bivariate distributions predicted from Propositions 1–6. Evidence for Propositions 7, 8, and 9, however, is somewhat mixed. Proposition 8 is directly supported by our data. Propositions 7 and 9 are supported for the partial effects of race. These effects are quite small for burglary and assault victimization, and the log-odds coefficients for personal larceny victimization, when rounded-off, equal zero. The clearest, seemingly disconfirmatory, evidence against Propositions 7 and 9 is found when one examines the effects of age on victimization risk. Contrary to our prediction, age was found to have consistent nonzero main effects on all three types of victimization at the multivariate level of analysis. Also, the three-way interactions involving age uncovered in our analysis were not anticipated by our theory. Once again, however, we propose that the existence of these interactions may be less of a challenge to our theory than they might at first appear. These effects may reflect the use of a proxy measure for the variables of exposure and guardianship. The propositions we deduced concerning the partial effects of age, income, and race presume that the effects of exposure and guardianship have been partialled-out. While we believe that the categories of the lifestyle variable we employ here do reflect systematic differences among persons with respect to degrees of exposure and guardianship, we also recognize that there may be significant variance within categories with respect to these two factors which may confound our analysis. Two examples of how age can be associated with variance within categories of exposure and guardianship are particularly relevant here.

First, by virtue of the facts that predatory offenders are disproportionately young and that age cohorts tend to be somewhat homogeneous in activity and interaction patterns, one would expect younger persons to have greater routine exposure to risk of predatory criminal victimization than older citizens. It is possible that the lifestyle variable employed herein better controls for homogamy by race and income as intervening factors than for homogamy by age. Of special relevance here is our control for unemployment status. Prior research suggests that the association of low income and race with unemployment (perhaps more so than low income or race per se) may account for the disproportionate tendency of offenders to be drawn from the low-income strata and to be nonwhite (see, for

example, Brenner, 1976). If this is indeed the case, then including unemployment status among the lifestyle categories employed in our study partials out exposure to risk due to homogamy by race and income. However, none of our lifestyle categories has the same direct relevance to homogamy by age, perhaps accounting for some unpredicted findings.

Second, we also expect that, within the categories of the lifestyle variable employed in the present study, increasing age is associated with more effective guardianship and decreased exposure to risk, by virtue of the fact that leisure activities away from home decline sharply around the age of twenty and continue to decline slowly as age increases (DeGrazia, 1961:121–4). For example, we expect that, although the unmarried and employed persons in our sample are, on the average, more likely to spend time outside of the home than some of our other lifestyle categories, the young within this category are more likely to do so than are older citizens. This increase in exposure to risk and decrease in guardianship effectiveness may be reflected in our findings.

The above discussion is also relevant to the three-way interaction effects involving age uncovered through our analyses. One speculation is that age, in combination with categories of lifestyle, income, and race, defines special categories of persons related to exposure that are not effectively tapped by our lifestyle variable. For example, the age-income combination identifies the young and poor as a category with higher victimization risk than would be expected on the basis of the main effects of age and income alone. One could plausibly interpret this interaction as identifying persons who are most likely to interact in time and space with high potential-offender subgroups, thus increasing exposure risk due to homogamy. We could offer an extensive list of such speculations here involving the combinations of age and other variables, but to do so leads us well beyond the data at hand. Instead, we emphasize that our findings regarding the effects of age underscore the need for more refined measures of dimensions of lifestyle as they relate to exposure and guardianship to be included in further victimization surveys if knowledge in this area is to advance.

CONCLUSION

In this paper, our efforts have been directed at formulating and systematically testing a theory of how certain dimensions of social stratification (i.e., income, race, and age) are related to the risk of predatory criminal victimization (e.g., assault, burglary, and personal larceny). In balance, the theory formalized and tested here appears to demonstrate considerable promise toward the end of developing a more general theory concerning the etiology of predatory victimization.

Our theory and data indicate that the relationship between the variables of income, race, and age, and the risk of predatory crime is a much more complex one than many criminologists realize. Many of our bivariate and multivariate findings are contrary to the conventional wisdom that "street crimes" (which include assault and personal larceny) occur disproportionately among low income persons (Carrington, 1975; Wilson, 1975), while household burglary increases with income (Scarr, 1972; Reppetto, 1974). Our data, however, indicate that, at the bivariate level, income is inversely related to risk of assault, directly related to risk of personal larceny, and parabolically related to risk of burglary victimization.

Also, at the bivariate level (and contrary to the conventional wisdom), we find race to have little direct effect on victimization risk for any of the crimes studied here. Age, on the other hand, is inversely related to each type of victimization at the bivariate level of analysis. When we control for proximity, guardianship, and exposure (introduced as factors associated with risk in our theory section), the income-victimization relationship changes. Most notably, the affluent have the highest risk of victimization for each crime studied here. The main effects for age and race uncovered in our log-linear analysis are similar to those observed at the bivariate level.

We have attempted to study offenses which are representative of Index Crimes

in general. Accordingly, we chose one violent crime (assault), one household crime (burglary), and one type of theft outside the home (personal larceny). On the basis of our findings, what can be said about how dimensions of social inequality (income, race, age) relate to these crimes?

First, age is related to victimization risk by virtue of its association with exposure, proximity, and guardianship patterns which we refer to here as lifestyle. The lifestyles of younger persons place them disproportionately in public places, at specific times, in closer proximity to potential offenders that increase the exposure of their person and property to risk of criminal victimization. Older citizens (particularly those 60 and older) less often expose their person or property to risk of victimization. Contrary to the considerable attention given by the media to victimization among the elderly, this age group has the lowest victimization risk among those studied here. However, the apparent low risk of victimization for those sixty and older may be deceptive. Given the relatively brief period of time each day that the average elderly person is exposed to risk, this group may be extremely vulnerable to victimization for such crimes as personal larceny and assault.

Our data further indicate that, independent of its relationship to the risk factors of lifestyle and proximity, income appears to have a direct effect on the probability of victimization for each of the crimes in our analysis. Given similar exposure to risk and lifestyle patterns, the affluent in our society apparently make the most attractive targets for crime, not the poor or middle-income citizens. Finally, our data suggests that, independent of its association with exposure, guardianship, and proximity, race has no appreciable effect on victimization risk. Nonwhites have just slightly higher risk of burglary and assault victimization than do whites, while whites have a slightly higher risk of personal larceny at the multivariate level of analysis.

Our interpretation of these findings is that social power resources relate to criminal victimization only in so far as they are collinear with differences in exposure to risk, guardianship patterns, proximity to potential offenders, and identification of lucrative targets. Independent of social power resources, however, differences in role-structured behavior, structural constraints, and routine vocational and leisure activities translate into differential victimization patterns among various groups in the U. S. population.

Some of our observed multivariate findings were not predicted from our theory. We have offered what we consider to be plausible hypotheses consistent with our theory that may account for the apparent discrepancies. The validity of these hypotheses as explanations for the variances observed between our theory and several multivariate findings await the development of more precise indicators of the risk factors identified here that intervene between our independent variables and predatory victimization. Despite the measurement problems inherent in these data (with respect to testing the theory presented herein), our current efforts have been encouraging and indicate that we are moving in the proper direction in our effort to predict and explain criminal victimization patterns in the United States.

REFERENCES

Brenner, Harvey M.
1976 Effects of the National Economy on Criminal Aggression II. Final report to National Institute of Mental Health. Contract number 282-76-0355FS.

Carrington, Frank G.
1975 The Victims. New York: Arlington.

Chapin, F.
1974 Human Activity Patterns in the City. New York: Wiley.

Cohen, Lawrence E. and David Cantor
1980 "The determinants of larceny: an empirical and theoretical study." Journal of Research in Crime and Delinquency 17:140–59.
1981 "Residential burglary victimization in the United States: lifestyle and demographic factors associated with the probability of victimization." Forthcoming: Journal of Research in Crime and Delinquency.

Cohen, Lawrence E. and Marcus Felson
1979 "Social change and crime rate trends: a routine activity approach." American Sociological Review 44:588–607.

Cohen, Lawrence E., Marcus Felson, and Kenneth C. Land
1980 "Property crime rates in the United States: a macrodynamic analysis, 1947–1977; with ex ante forecasts for the mid–1980s." American Journal of Sociology 86:90–118.

Cooley, Charles H.
1909 Social Organization. New York: Scribners.
DeGrazia, S.
1961 "The uses of time." In R. Kleemeier (ed.), Aging and Leisure. New York: Oxford University Press.
Farley, R.
1977 "Residential segregation in urbanized areas of the United States in 1970: an analysis of social class and racial differences." Demography 14:497–518.
Garofalo, James
1979 "Victimization and the fear of crime." Journal of Research in Crime and Delinquency 16:80–97.
Goodman, L. A.
1972 "A general model for the analysis of surveys." American Journal of Sociology 77:1035–86.
1973 "Causal analysis of data from panel studies and other kinds of surveys." American Journal of Sociology 78:1135–91.
1976 "The relationship between modified and usual regression approaches to the analysis of dichotomous variables." Pp. 83–110 in D. R. Heise (ed.), Sociological Methodology 1976. San Francisco: Jossey-Bass.
Hindelang, Michael J.
1975 Public Opinion on Crime, Criminal Justice, and Related Topics. U. S. Department of Justice, Law Enforcement Assistance Administration, National Criminal Justice Information, and Statistics Service. Analytic Report 1. Washington, D.C.: U. S. Government Printing Office.
1978 "Race and involvement in crime." American Sociological Review 93:93–109.
Hindelang, Michael S., Michael R. Gottfredson, and James Garofalo
1978 Victims of Personal Crime. Cambridge: Ballinger.
Land, Kenneth C., and Marcus Felson
1976 "A general framework for building dynamic macrosocial indicator models: including an analysis of changes in crime rates and police expenditures." American Journal of Sociology 82:565–604.
Nettler, Gwynn
1978 Explaining Crime. New York: McGraw.
1979 "Criminal justice." Annual Review of Sociology 5:27–52.
Olsson, Gunnar
1964 Distance and Human Interaction: A Review and Bibliography. Philadelphia: Regional Science Research Institute.
Penick, B. and M. E. B. Owens
1976 Surveying Crime. Washington D.C.: National Academy of Sciences.
Quinney, Richard
1975 Criminology: An Analysis and Critique of Crime in America. Boston: Little Brown.
Reppetto, Thomas J.
1974 Residential Crime. Cambridge: Ballinger.
Riley, Matilda, Marilyn Johnson, and Anne Foner
1972 A Sociology of Age Stratification. Volume 3 of Aging and Society. New York: Russell Sage Foundation.
Scarr, Harry A.
1972 Patterns of Burglary. U. S. Department of Justice, Law Enforcement Assistance Administration. Washington, D.C.: U. S. Government Printing Office
Taueber, Karl E. and A. F. Taueber
1965 Negroes in Cities. Chicago: Aldine.
U. S. Bureau of the Census
1973 "Who's home when?" Washington, D.C.: U. S. Government Printing Office.
1979 Marital Status and Living Arrangements: March 1978. Current Population Reports. Series P-20, No. 338. Washington, D.C.: U. S. Government Printing Office.
U. S. Department of Justice
1974a Preliminary Report of the Impact Cities. Crime Survey Results. Washington, D.C.: Law Enforcement Assistance Administration.
1974b Crime in the Nation's Five Largest Cities: Advance Report. Washington, D.C.: Law Enforcement Assistance Administration.
1974c Crimes and Victims: A Report on the Dayton–San Jose Pilot Survey of Victimization. Washington, D.C.: Law Enforcement Assistance Administration
1976 Criminal Victimizations in the U. S., 1973. Washington, D.C.: Law Enforcement Assistance Administration.
1977 Criminal Victimizations in the U. S.: Comparison of 1974 and 1975 Findings. Washington, D.C.: Law Enforcement Assistance Administration.
Walsh, Marilyn
1974 "A study of the fence and how he operates." Doctoral dissertation. Ann Arbor, Michigan: University Microfilms.
Wilson, James O.
1975 Thinking About Crime. New York: Basic.

SOCIAL CONTROL THEORY AND DELINQUENCY*

MICHAEL D. WIATROWSKI
DAVID B. GRISWOLD
Florida Atlantic University
MARY K. ROBERTS
University of Florida

Hirschi's social control theory proposes that delinquents fail to form or maintain a bond to society consisting of attachment, commitment, involvement, and belief. Using data from the Youth in Transition Study, the present report develops and tests multivariate models of social control theory which simultaneously consider how the four bond elements operate in relation to delinquency. Factor analysis and communality analysis examine the uniqueness of the four bond elements, and revised and additional measures are suggested. Background factors—measures of social class and ability—are added to the model, and a revised formulation of social control is suggested.

Hirschi's (1969) *Causes of Delinquency* is a benchmark for theory construction and research in the delinquency field. The theory rests on the Hobbesian assumption that human behavior is not inherently conforming, "but that we are all animals and thus naturally capable of committing criminal acts" (Hirschi, 1969:31). Since delinquency is intrinsic to human nature, it is conformity that must be explained. Conformity is achieved through socialization, the formation of a bond between individual and society comprised of four major elements: attachment, commitment, involvement, and belief. The stronger each element of the social bond, the less likely delinquent behavior.

* Direct all correspondence to Michael D. Wiatrowski, Department of Criminal Justice, Florida Atlantic University, Boca Raton, FL 33431.

Special thanks to Dr. Gary Gottfredson and Dr. Don C. Gibbons for their assistance with this manuscript. Thanks also to Dr. Jerald Bachman, Dr. Travis Hirschi, and the reviewers for their comments on earlier drafts.

The data used, originally collected by Jerald Bachman, were made available by the Inter-University Consortium for Political and Social Research. The responsibility for the analyses or interpretations presented here is mine alone. Preparation of this report was supported by grants NIE-G-78-0210 and NIE-G-80-0113 to the Center for Social Organization of Schools, The Johns Hopkins University. The opinions expressed in this report do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the Institute should be inferred.

Attachment corresponds to the affective ties which the youth forms to significant others. The family environment is the source of attachment because parents act as role models and teach their children socially acceptable behavior.

Commitment is related to the aspiration of going to college and attaining a high-status job. This is an investment in conventional behavior which the youth risks should he become delinquent. In contrast to youths with well-defined goals, adolescents engaged in drinking, smoking, dating, and other behavior not oriented toward future goals are much more likely to get involved in delinquent behavior.

Involvement refers to participation in conventional activities which lead toward socially valued success and status objectives. The quality of a youth's activities and their relationship to future goals and objectives are important in preventing delinquency. Time spent on homework, for example, is viewed as antecedent to success in attaining educational goals which are prerequisites to high-status occupations.

Belief is acceptance of the moral validity of the central social-value system (Hirschi, 1969:203). This variation in the acceptance of social rules is central to social control theory, because the less rule-bound people feel, the more likely they are to break rules (Hirschi, 1969:26). Hirschi (1969:26) argues that there is one

dominant set of values and that even delinquents may recognize the validity of those values, although they may not feel bound by them because of weakened ties to the dominant social order.

HIRSCHI'S DATA ANALYSIS

Although Hirschi's theory fares better than subcultural and differential association theory, Empey (1978:239) implies that the theory's empirical support falls short of complete explanation. Hirschi does not consider how his four elements might act simultaneously to affect the likelihood of delinquent behavior. In addition, instead of empirically analyzing the relationships among the elements of the bond, he simply hypothesizes relationships between attachment and commitment, commitment and involvement, and attachment and belief. Consequently, Hirschi's theory construction and data analysis raise three related questions. First, the extent to which Hirschi's four elements represent empirically distinct components of socialization is unclear. If most of the variance explained in the criterion is shared by the four elements, they would not constitute analytically distinct elements of the bond. Second, why are only four elements of the bond identified? The modest predictive power of Hirschi's constructs suggests that additional elements of the bond should be considered. Third, although educational and occupational aspirations are central to Hirschi's theory, he fails to incorporate constructs—such as family socioeconomic level, ability, and significant-others' influence—that research has determined to be important in the development of these aspirations (Haller and Portes, 1973; Sewell, Haller, and Portes, 1969). Hirschi worries about this, but his examination of the zero-order correlations of delinquency and social class finds that there is no important relationship between social class and delinquency (1969:75). Nevertheless, Hirschi (1969:73) suggests that this relationship may be suppressed by some third, intervening variable.

The present research addresses the foregoing three issues. First, we create measures of each of Hirschi's four elements and estimate how much each element contributes over and above the others to the explanation of delinquent behavior. Second, we examine more closely the structure of the social bond, using factor analysis to discover and define new elements. Finally, we perform two multivariate analyses, incorporating measures of social class, ability, and grades as well as elements of the social bond. The first analysis, using Hirschi's four elements, is intended to replicate and extend, rather than extensively reformulate, his study. Finally, we develop and evaluate a revised model, based on the factor analysis, which is somewhat more parsimonious and complete than Hirschi's.

DATA AND MEASURES

Data were obtained from the Youth in Transition Study (Bachman, 1975). This is a longitudinal study of 2213 tenth-grade boys from 87 schools who were surveyed in 1966 in five waves of data collection. The use of concurrent measures parallels Hirschi's (1969) research. A multistage probability sample was used to provide an essentially unbiased representation of tenth-grade boys in public high schools throughout the United States (Bachman, O'Malley, and Johnston, 1978:3).[1]

The selection of information was guided by Hirschi's (1969) discussion and research. Since the exact items used in Hirschi's study were not available, equally weighted composite scales were constructed from items which coincided with Hirschi's constructs. The selection of dissimilar items creates the possibility that differences in results are attributable to

---

[1] Originally 88 schools selected from the University of Michigan's Survey Research Center's primary sampling units were invited to participate in the study and 71 consented. Replacement schools were found in the same geographic area for all but one school. Approximately 25 males were selected from each school, and of the 2,277 boys who were invited to be members of the survey group, 2,213 (over 97%) chose to participate. This would indicate that nonrespondents and dropouts were not a serious problem (Bachman, O'Malley, and Johnston, 1978: 2–4). Participation rates for waves 2 through 5 were 85.2, 81.3, 73.2, and 73.5%.

differences in measures. Homogeneity or reliability coefficients are used to assess scale reliability. Although assessments of acceptable levels of reliability are governed by convention, a low coefficient is a source of measurement error. In the more parsimonious model of the social bond, selected items are deleted from scales when they result in scales with unacceptably low reliability coefficients.

Items in the Youth in Transition data set that appeared closely related to those in Hirschi's data set were identified. Then factor-analytic and item-analytic procedures were used to identify the measures to be used in the model testing. The following paragraphs discuss the items chosen for initial examination as measures of Hirschi's four elements; see Appendix A for abbreviations and sample items.

Attachment involves the relation of the youth to parents, peers, and school. We represented attachment to parents by two indices, measures of closeness of the youth to his father and to his mother. Attachment to peers was represented by items about the importance of friends to the respondent and how important it was to spend time with his peers. Attachment to school was represented by indices of positive attitude toward school, negative attitude toward school, academic achievement, self-concept of school ability, and the youth's perception of teachers' interest in him.

Commitment was represented by an index of the youth's occupational aspirations coded by Duncan occupational prestige scores. Educational aspirations were measured by a continuum of educational interests; additional items measured the clarity of occupational plans, whether the boys had taken any steps toward attending college, and the amount of time spent and frequency of dating. As will be shown later, the wide conceptual range in the component elements of commitment, as initially constructed, affected its reliability coefficient.

Involvement was represented by three pieces of information that indicated how much school work the youth was doing. While Hirschi relied primarily on the relationship between the amount of homework done and delinquency, we expanded the analysis to include not only homework but also extra school work not required by the teacher and the frequency with which school work was discussed by the boys.

Belief was difficult to operationalize because the process of developing beliefs is complex. A youth with close parental attachments is rewarded for conformity by ". . . the approval and esteem of those he admires" (Hirschi, 1969:200). This causal sequence in turn leads to a respect for ". . . persons in positions of authority, to belief that the rules of society are binding on one's conduct" (1969:203). Frequently, Hirschi uses "conscience" in terms of superego development (1969:87) and integrates this with the concept of guilt in noting that delinquent youth experience little guilt when breaking the law.

We used a broader definition of belief, selecting items that measured belief at its final stage of development rather than some earlier stage. The possibility does exist that youth may abide by social rules but not respect the police. An honesty index comprised of items which assessed attitudes towards truthfulness, lying, cheating, and helping friends in difficult situations was used for part of the belief measure. These belief items were initially supplemented by a guilt index, included because individuals lacking a psychological sense of responsibility would, in theory, be free to behave without any notion of psychological accountability for their behavior. The guilt index was later deleted because of unacceptable psychometric properties of the belief scale; its relatively low reliability coefficient is a potential challenge to the use of this index as an operational measure of belief.

Delinquency was measured using an index consisting of the responses to 26 items (with up to six missing data items allowed) adapted from Gold's (1966) self-report measure of delinquency. Among the areas measured were responses to questions about theft and vandalism, interpersonal aggression, delinquency in school, frequency and seriousness of delinquency, and trouble with parents. To protect the confidentiality of the respondents, data on individual delinquency

items were not made available. While it would have been desirable to restrict the items to criminal offenses and exclude trivial offenses, this was not possible. However, the items included in the index have a high alpha coefficient of .85, which demonstrates their internal consistency.

Hirschi's study used two items from Nye and Short's (1957) seven-item delinquency scale and four items from Dentler and Monroe's (1961) five-item "theft scale." Nye and Short's research has been criticized for including trivial items (Hindelang et al., 1979) as have other early self-report instruments (Elliott and Ageton, 1979). Differences in the findings of this study and those of Hirschi may be due to the different measures of the dependent variables.

In the Youth in Transition study report there was no direct validation of delinquency measures against official police reports. There is evidence, however, concerning the validity of the scale items (Bachman, et al., 1978:172) and the self-report method in general (Elliott and Ageton, 1980:96; Hardt and Hardt, 1977). These include the concurrence of the Youth in Transition findings with other studies, the stability of responses over an eight-year period, the internal consistency of attitudinal and behavioral data, and methodological studies that indicate infrequent affirmative responses to the use of fictitious drugs. The authors conclude that the validity of a variable is a matter of individual judgment and that their ". . . measures are basically valid" (Bachman, O'Malley, and Johnston, 1978:173).

Socioeconomic level was measured with a scale of five items that were equally weighted to form an index: father's occupational status, parent's education, possessions in the home, number of books in the home, and the number of rooms per person in the home. Mental ability was measured using scores on the General Aptitude Test Battery for verbal and math ability.

COMMUNALITY ANALYSIS

Hirschi theorizes that there are four separate elements of the social bond. In statistical terms this means that the four measures are independent of each other—that each makes at least some unique contribution to predicting delinquency. The extent to which the elements do make unique contributions is examined in this section. First, the extent to which delinquency is predictable using all the measures assumed to tap some element of the bond is estimated. Then composite measures, constructed on the basis of Hirschi's (1969) theory, are used to predict delinquency, and the unique contribution of each composite is examined.

Tables 1 and 2 summarize the first part of these results. The proportion of the variance of delinquency explained by all 23 individual indicators of the bond was .318. The unique variance attributable to each element of the bond was obtained by subtracting the squared multiple correlation of all bond elements except those assumed related to the element under con-

Table 1. Regression of Delinquency on Individual Measures Assumed to be Associated with the Elements of the Bond

Bond element and variable	r	Beta
Attachment		
CLOSMOM	−.261*	−.120*
CLOSFATH	−.243*	−.101*
HOWIMPF	−.015	.011
TIMWFRN	.015	−.006
POSSCHI	−.313*	−.083*
NEGSCHI	.302*	.104*
ACAACHI	−.256*	−.028
ABILCON	−.102*	.028
TCHINTR	−.163*	−.045**
Commitment		
DUNASPI	−.085*	.060**
CLAROCPL	.015	−.016
RCVOJT	.039	−.016
COMPHS	−.137*	−.033
RCVMILT	.036	.015
RCVVOC	.021	−.020
ATNDCOL	−.158*	−.028
MADECOLP	.018	.013
DATEIND	.372*	.332*
Involve		
TIMEHW	−.155*	−.082*
DSCHWFR	−.160*	−.012
XTRASCH	−.181*	−.050**
Belief		
HONESTI	−.319*	−.136*
GUILTIN	−.155*	.069*
R^2 = .318		

* $p \leq .05$.
** $p \leq .01$.

Table 2. Total Association and Unique Contribution of Each Category of Predictors

Set of predictors	Number of measures in the set	Total association	Unique contribution
Attachment	9	.177	.065
Commitment	9	.157	.109
Involvement	3	.061	.010
Belief	2	.103	.016
All Predictors	23	.318	—

NOTE: Total association is the squared multiple correlation of a set of predictors with delinquency. The unique contribution of a set of predictors is the incremental validity of the set. That is, it is the gain in R^2 achievement when that set is added to the regression equation after all other predictors have already been used in a regression equation.

sideration from the squared multiple correlation obtained when all bond elements are used to predict delinquency. In Table 3 the variances for each element of the bond are indicated along with the explained variance for an element operating by itself. The unique variances for each element are quite low, ranging from .010 for involvement items to .109 for commitment. An alternative interpretation of the low amount of uniqueness is that underlying these measures of the bond there exists a general factor called socialization, and that these elements along with other undefined elements of the bond are negatively correlated with delinquency because delinquency is negatively correlated with the larger construct of socialization.

One criticism of the communality analysis presented above is that categories of bond measures containing a larger number of measures may be expected to be associated with more variance in the criterion because of their number alone. In addition, the use of multiple indicators of each bond element does not allow for a simple presentation of control theory. To deal with these problems, scales were formed for each bond element and the communality analysis was again performed.

Scales were constructed by examining the correlation matrix and determining which items within an element such as attachment or commitment were positively correlated with each other. Appendix B describes the measures used in Tables 3, 4, 5, and 6. The correlation matrix and alpha reliability coefficients for each scale are presented in Table 3. The commitment and belief scales were difficult to construct because the items which were initially chosen to operationally represent an element formed scales with unacceptably low alphas. Only scales with acceptable reliabilities are reported here. For the commitment scale the dating item did not form a scale with the occupational and educational aspiration items. The proportion of variance of delinquency involvement explained by a scale formed from all three items was less than that explained by dating alone. One interpretation of this is that involvement in dating may represent a dimension of the bond which is independent of the other parts of commitment, as the low communality of dating in the factor analysis in the following section would indicate.

The honesty and guilt indices formed a scale with an alpha of .22. Despite the allusions to the relationship of conscience and superego development to belief (Hirschi, 1969:87, 200), the data did not support combining these measures into a

Table 3. Correlation Matrix and Alpha Coefficients for Scaled Bond Measures

	1	2	3	4	5	Alpha[a]
1. Attachment	1					.54
2. Commitment	.321*	1				.59
3. Involvement	.403*	.205*	1			.77
4. Belief	.435*	.193*	.192*	1		.87
5. Delinquency	−.360*	−.137*	−.248*	−.319*		.85

* $p \leq .01$.

[a] This is an approximation based on code book data for item means and standard deviations, and the total scale mean and standard deviation, using an adaptation of the formula for KR20. Estimates of the reliability of several subscales made by Patrick O'Malley (personal communication, August 30, 1979), assuming that measurement error is equal at each time (i.e., each data collection) and that errors are uncorrelated, range from .85 for a scale composed of items related to delinquent behavior in school to the low .50s for other scales.

simple index. The correlation between honesty and guilt was .129 and the coefficient between guilt and delinquency was −.015. The guilt item was *deleted* from the analysis and the homogeneity coefficient of .87 shown for belief in Table 3 refers to the reliability of the honesty index. The zero-order correlation between the honesty measure and delinquency is −.319, in the direction and magnitude predicted by Hirschi.

In Table 3 the correlation matrix shows that the scaled bond elements are associated with each other and that each element is negatively related to delinquency to about the same degree as was reported by Hirschi (1969). The proportion of total explained variance of the four bond elements taken together as shown in Table 4 declined to .174 because of the exclusion of the dating item from the commitment scale.[2] The unique variance explained by each scaled element remained small, ranging from .000 to .034, indicating that incrementally each bond element adds little to the explanation of delinquency in relation to the total amount of variance explained.

Table 4. Squared Zero-Order Correlations and Unique Contributions of Scaled Elements of the Bond

Element	r^2	Unique variance
Attachment	.130*	.034
Commitment	.019*	.000
Involvement	.062*	.012
Belief	.102*	.032
R^2	.174	

* $p \leq .01$

INTERNAL STRUCTURE OF THE BOND

Our variables were chosen as reasonable measures of Hirschi's constructs, and considerable efforts were taken to examine the psychometric properties of these variables. The items assumed to represent elements of the social bond were factor analyzed using a principle component analysis and varimax factor rotation to examine their underlying structure. We would expect that four elements representing attachment, commitment, involvement, and belief would emerge as factors. The structure of the bond, however, takes a different form from that implied in *Causes of Delinquency.*

Table 5 shows the results of a seven-factor principal components solution with varimax rotation. Although a scree test implied that fewer factors might have been rotated, the conceptual clarity of the seven-factor solution led us to use it for this portion of the analysis (seven eigenvalues were greater than 1.0).

The positive and negative school indices and the academic achievement index have high loadings on the first factor (Table 5), which might be termed attachment to school. The second and third factors represent status or achievement orientation and appear related to Hirschi's concept of commitment. Those individuals who have high aspirations, are certain of their academic abilities, and want to attend college score high on factor II. Correspondingly, the vocational items including job training and military or vocational training are negatively loaded on this factor. Boys scoring high on the third factor have low academic-ability self-concepts, low aspirations, and unclear occupational plans. Those represented on this low-status orientation factor are also unlikely to expect to complete high school or attend college.

The fourth factor appears to tap school involvement, including the positive school-attitudes index and the feeling that teachers take a personal interest in that individual. This dimension also taps a "motivational" element, in which the youth is willing to discuss homework with friends and voluntarily do extra school work.

The final three factors explain only a small portion of the remaining variance. Factor V relates to parental attachment. Factor VI suggests the existence of a peer-attachment element of the social bond. The dating element also appears on factor VI, suggesting that this dimension

[2] The association of extensive dating with a host of youthful behaviors which do not conform to conventional adult values is well-supported in the literature (see Hirschi, 1969:163–71; Galvin, 1975; and Coleman, 1961). However, Bealer, Willits, and Maida (1965) have argued that failure to adopt one or more conventional values does not preclude conformity to others.

Table 5. Varimax Rotated Factor Structure of Measures Related to Hirschi's Constructs

	I	II	III	IV	V	VI	VII	h^2
CLOSMOM	.207	−.008	−.006	−.088	.771	.061	.042	.650
CLOSFATH	.068	−.023	−.033	−.160	.805	−.007	.001	.679
HOWIMPF	.054	.076	−.027	−.016	.073	.726	.024	.543
TIMWFRN	.068	−.023	.036	−.008	−.049	.777	−.005	.612
POSSCHI	.565	.003	−.173	.471	.198	.090	.097	.628
NEGSCHI	−.550	−.213	.303	−.203	−.167	−.041	.119	.524
ACAACHI	.748	.055	−.141	.149	.057	.057	.121	.626
ABILCON	.132	.443	−.354	.328	−.080	.041	.135	.473
TCHINTR	.028	.019	.016	.568	.273	.128	−.077	.421
DUNASPI	.215	.419	−.481	.112	−.092	.054	.070	.482
CLAROCPL	.021	.161	−.505	−.044	.057	−.030	−.108	.299
RCVOJT	−.057	−.720	−.052	.077	−.036	.051	−.001	.534
COMPHS	.114	.014	−.689	−.031	.059	.041	−.046	.496
RCVMILT	.026	−.659	.043	.038	−.008	.004	.012	.438
RCVVOC	−.082	−.682	−.044	−.008	.039	.002	.017	.476
ATNDCOL	.153	.443	−.572	−.147	.043	.037	.088	.580
MADECOLP	−.126	.009	−.176	−.033	.107	.006	.802	.703
DATEIND	−.332	−.191	.019	.016	.071	.365	.030	.286
TIMEHW	−.053	−.001	.277	−.190	−.086	.091	−.057	.134
DSCHWFR	.114	−.033	−.046	.701	.004	.024	−.004	.508
XTRASCH	.103	.064	−.050	.737	.039	−.094	.064	.575
HONESTI	.799	.026	−.067	.010	.136	−.107	−.001	.662
GUILTIN	.328	.015	.177	.076	−.091	.054	.575	.487

NOTE: Complete names for variables and sample items included in indexes are provided in Appendix A.

taps a more comprehensive concept of sociability in which the youth chooses to associate not only with his male peers but also with members of the opposite sex. Factor VII is difficult to interpret and represents little of the common variance. The only item with a high loading indicates whether the youth had made college plans. This item was chosen to show whether the youth was able to perceive the link between educational aspirations and actually attending college. It would have been expected to load on the second factor, and the failure to do so may be interpreted as an indication that attitudes and behavior are not always linked, or that too many factors have been extracted.

Because the four elements of the social bond representing attachment, commitment, involvement, and belief did not appear as separate factors, a new interpretation of the bond appears necessary. We find factors representing parts or components of the social bond, such as the attachment to school and school involvement, high and low status commitments, and parental and peer attachments. The dating and belief items also did not appear as factors. Dating had a moderate negative loading on attachment to school and the high-status career orientation but was unrelated to the low-status career stream. This is important because a host of nonproductive juvenile activities which include dating, drinking, and cruising around in a car are thought to prevent youths from making investments in conventional behavior. While this is partially supported in these data, concomitantly, dating is largely unrelated to all of the factors as a whole ($h^2 = .29$). The low communality of dating suggests that it should be represented separately in a social-control model.

These results suggest a more complex interpretation of the social bond than that presented by Hirschi. The emergence of a strong factor involving the school accords with other research which indicates that in adolescence the peer structure of boys is a major locus of influence (Greenberg, 1977; Coleman, 1961; Smelser and Halpern, 1978; Polk and Schafer, 1972). The presence of this school-related factor is also consistent with the view that one function of the school is to assist young people in the transition to adult social roles. School serves as a mechanism in which aspirations formed earlier in life are translated

into concrete achievement goals. Naturally, conclusions derived from any factor analysis are largely determined by which variables are chosen for analysis. While Hirschi's analysis reflects a concern with parental, educational, peer, belief, and aspirational items, it is grounded on the premise that the parental relationship is an important determinant of the later elements of the social bond. Thus the emergence of school-related factors is in part a function of the number of school-related items selected, and it is also indicative of the relative strength of those items in a factor analysis of the social bond.

A factor representing a vocational orientation implies that there is a group of youths bonded to society, but in somewhat lower status positions. This factor appears to accord with some speculation by Polk (1975) that the relationship between social status and socialization includes lower social status youths who are not involved in an alternative youth culture system. The zero-order correlations of dating with a scaled measure of commitment and vocational orientation (described in a later section of the report) are close to zero (.000 and .085, respectively). These results do not demonstrate the ability of commitment to conventional goals to exclude dating or the preoccupation of those vocationally oriented with activities which are unrelated to future goals.

The emergence of the honesty index on the school-attachment dimension contradicts Hirschi, who hypothesized that the adherence to conventional social values should be related to attachment to parents. This is consistent with research which suggests that, as society becomes more complex, socialization functions which once belonged to the family are assumed by educational institutions (Smelser and Halpern, 1978; Parsons, 1959).

The variety of items in the factor analysis loading on the first factor (attachment to school) indicates that youths with a positive relationship to school are making investments in conventional patterns of behavior. This is congruent with the thesis that school does have a socializing function in which values are reinforced and also with a social control hypothesis that school involvement represents a primary group process in which socialization occurs in successful, conventional social interaction.

In summary, some of Hirschi's postulated dimensions emerge as distinct factors, but the general picture of the components of the bond is altered. What this suggests is that it may be more appropriate to reconceptualize the nature of the bond.

A SIMPLE MULTIVARIATE MODEL OF THE SOCIAL BOND

We constructed a path model which structured the multivariate examination of Hirschi's four elements of the bond (Figure 1), despite doubts about the elements' utility raised by the communality and factor analyses. Rather than simply arguing, as did Hirschi (1969:75), that social class is not important to delinquency, we consider socioeconomic class and ability as prior or exogenous variables whose causes are unanalyzed, but examine direct and indirect effects of these variables via elements of the bond. The contributions of the four elements of social control theory are also assessed with SES and ability used as statistical controls. Thus this "simple" model may be considered a modification of Hirschi's original formulation. However, if the hypothesized effects do not emerge, then notions of class and ability differences in socialization can be discarded and Hirschi's theory returned to its original form. Appendix B describes the indices which are analyzed in Tables 6, 7, 8, and 9.

Table 6 shows the decomposition of effects according to the model in Figure 1, and Table 7 shows the direct path coefficients according to the model.[3] The re-

[3] Path analysis was used to examine the effects of components of the model described in Figure 1 on delinquency outcomes. Central to our analysis is an understanding of the terms total associations, total contributions, direct contribution, and indirect contributions (Alwin and Hauser, 1975; Gottfredson, 1978). The total association is the zero order correlation between two variables, while the total contribution is the standardized partial regression coefficient (Beta weight) in a regression equation which includes all potential explanatory variables. The direct con-

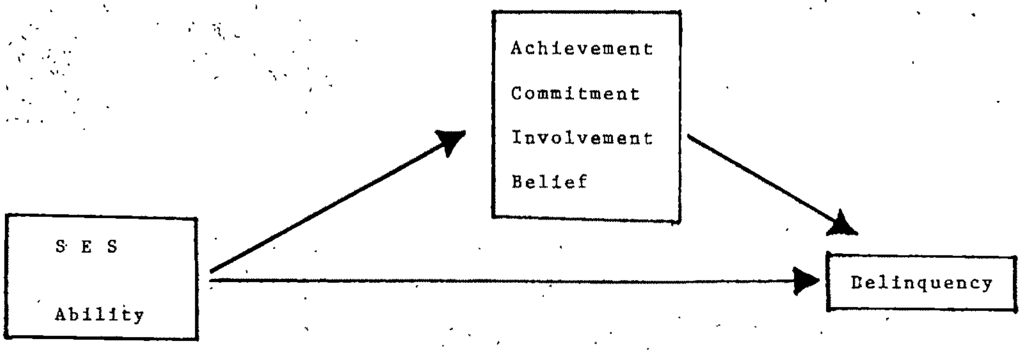


Figure 1. Simple Multivariate Model of the Social Bond

Table 6. Decomposition of Effects According to the Path Model Involving Social Control Theory Elements of the Bond

Independent variable	Total association	Total effect	Direct effect
Background			
SES	.026	.044	.103*
Ability	−.024	−.043	−.060**
Bond			
Attachment	−.360*	−.234*	−.234*
Commitment	−.137*	−.069*	−.069**
Involvement	−.248*	−.113*	−.113*
Belief	−.319*	−.206*	−.206*

NOTE: R^2 = .188 residual = .901
* $p \leq .01$
** $p \leq .05$

sults imply that each of the social control theory variables makes significant direct contributions to the explanation of delinquent behavior even when SES and ability are statistically controlled, and that none of the contributions is negligible, even when other social control variables are considered simultaneously. The direct and indirect contributions of the background variables (SES and ability) are of interest. As noted by Hirschi (1969), social class has a tiny and nonsignificant total association with delinquency. When other variables are considered simultaneously, however, SES paradoxically makes a modest but significant positive contribution to the amount of self-reported delinquent behavior. Hartshorne and May's *Studies in Deceit* (1928) provides important insight into this relationship by documenting the association between honesty and intelligence. Youth who are from higher status backgrounds and are more intelligent may report their deviant behavior more completely, accounting for the relationship between socioeconomic status, intelligence, and delinquency.[4]

Social class is a significant factor in explaining the levels of commitment and of moderately significant importance in explaining levels of attachment and involvement. Ability also makes moderately sized and significant positive contributions to attachment and belief, and it has a major effect on the level of commitment. This outcome accords well with Hirschi's theoretical account, in which ability is assumed to influence levels of delinquent behavior primarily because it affects the commitment of youth to conventional attainment goals. It should be kept in mind that commitment as measured in the present research refers largely to high levels of academic and occupational aspirations. In short, the results correspond with theory in implying that students of low ability and social class are less committed to conventional goals, presumably because those goals are beyond their reach, and that as a consequence of lowered commitment, these students are free to engage in delinquent behavior.

In spite of the generally positive results, these findings are far from satisfactory for three reasons. First, the proportion of

tribution is the Beta weight in a regression equation which includes the causally prior but not intervening variables at that stage, or inclusion level (cf. Nie et al., 1975:375) of the equation, while the indirect effects are those effects which are transmitted via intermediary variables.

[4] The book by Hartshorne and May, *Studies in Deceit* (1928) was made known to the author by an anonymous reviewer.

Table 7. Standardized and Unstandardized Path Coefficients in Social Control Model of Delinquency

Independent variable	Attachment Beta (b)	Commitment Beta (b)	Involvement Beta (b)	Belief Beta (b)	Delinquency Beta (b)
Background					
SES	.122(.006)*	.287(.006)*	.092(.002)*	.006(.000)	.103(.000)*
Ability	.149(.058)*	.315(.057)*	.017(.004)	.214(.023)*	−.060(.002)**
Bond					
Attachment					−.234(−.019)*
Commitment					−.069(−.012)**
Involvement					−.113(−.016)*
Belief					−.206(−.059)*
R^2	.053	.260	.010	.047	.188
Residual	.973	.860	.999	.976	.901

* $p \leq .01$.
** $p \leq .05$.

variance in delinquency explained by the model is not large—19%. This is small in comparison with the proportion explained using all twenty-three variables examined in Table 5, primarily because the predictive power of the individual variables was ravaged by constructing scales which accorded closely with Hirschi's theoretical statement. In particular, because dating did not scale with the commitment variables as Hirschi appears to imply it would, this variable could not be used. (Scoring it together with other measures of commitment results in a much lower reliability of that scale. The alpha reliability of the scale used was .59, and when dating is added this drops to .46.) Second, the factor-analysis results imply that an alternative set of bond elements would more faithfully represent the structure of the variables involved. And third, the model does not explicitly take into account the well-established finding that school grades are inversely related to delinquency (Hirschi, 1969:111–20; Silberberg and Silberberg, 1971; Bachman, O'Malley, and Johnston, 1978), which implies that the model is misspecified. The next section describes analysis of a reformulated model designed to remedy these defects.

A COMPLEX MODEL OF THE SOCIAL BOND

In the reformulated model (Figure 2) the bond elements are chosen to more faithfully represent the bond components and structure derived from interpretation of the factor and communality analyses.[5] Since Hirschi did not present a multivariate model of the social bond, the model here can be challenged on the grounds that the specification of the model is incorrect. However, this should indicate the general difficulty of constructing models from verbal descriptions of theories. Socioeconomic status and ability are again treated as exogenous background variables because status-attainment research (Blau and Duncan, 1967; Haller and Portes, 1973) implies that SES and ability affect the nature of parental socialization, which in turn affects educational and occupational aspirations and attachment to school. The relationship of ability to educational aspirations, school attachment, and grades through parental attachment is also informed by the status-attainment model. Parental attachment was regarded as the foundation of the social bond. Thus, the model shows parental attachment as causally prior to and directly linked with commitment to educational and occupational aspirations, dating, attachment to school, and involvement. While factor analysis says nothing about the causal ordering of variables, belief was placed after the previous block of variables because of the loading of the belief items on the attachment-to-school factor. This suggests that acceptance of conventional social values may be the consequence of a youth's belief in the efficacy of education in pursuing future goals, although the relationship could

[5] The measures used in the remainder of this analysis are presented in the bottom portion of Table

5. Dr. Travis Hirschi reviewed an early draft of this paper and did not criticize the specification of the model.

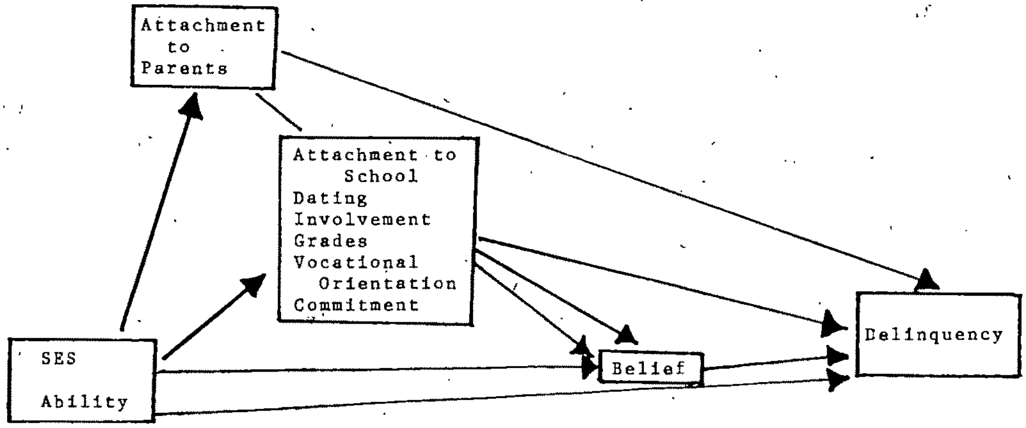


Figure 2. Complex Model of the Social Bond

be reversed. Belief is treated as a separate element because it is conceptually distinct from attachment to school.

The revised model of social control theory explains 32.5% of the variance in delinquency. With 13 fewer variables, the amount of variance in delinquency explained is comparable to that explained in the 23-item regression analysis presented earlier in Tables 1 and 2. Compared to the simple social control model shown in Figure 1, the more complex model of Figure 2 explains an additional 14% of the variance.

Table 8 shows the decomposition of effects according to the revised model and Table 9 shows the path coefficients. Parental attachment and school attachment both have substantial negative total and direct effects on delinquency, net of other variables in the model; and dating has substantial positive total and direct effects. Grades, involvement, and belief have smaller but nevertheless statistically significant total and direct effects on delinquency. In contrast to the earlier model (Figure 1), the revised model implies moderate positive *net* (direct) effects of SES and ability. Grades make a moderate negative direct contribution, as does involvement. In the context of the revised model, commitment to a vocational career, commitment to college and a high-status occupation, and peer attachment make no statistically significant contributions to the explanation of delinquency, implying that their zero-order association may be regarded as spurious.

In social control theory, the peer relations of delinquents are characterized by their low level of social skills (Hirschi, 1969). The lack of a relationship between peer attachment and delinquent behavior therefore reflects the unimportance of friends to delinquent youth. This illustrates an important difference between social control theory and other delinquency theories which posit a central role for peer relations (Hansell and Wiatrowski, 1980).

Some of the other results in Tables 8 and 9 should be highlighted. First, except for belief and grades in general, the coefficients of determination (R^2) are generally small, implying that the residuals (the contributions of unmeasured variables and measurement error) are large. This means, in short, that even the revised model allows much room for improvement. Second, ability has strong positive total and direct effects on grades, and moderate positive total and direct effects on school attachment, and a moderate negative effect on dating. Parental attachment has fairly strong positive effects on school attachment, involvement, and belief (as well as delinquency), implying that even for adolescents who are well into their high-school years, parental attachment exerts considerable influence.

The proportion of variance in belief which was explained is substantial in comparison to other elements of the bond. This analysis makes possible a comparison of the effects of parental versus school attachment on belief. The path coefficient

Table 8. Decomposition of Effects of Revised Model of Social Control Theory

	Total association	PARATT		GRADETM1		DATEIND		SCHATT		VOCORNT	
		Total	Dir.	Total	Dir.	Total	Dir.	Total	Dir.	Total	Dir.
SES	.025	.062*	.062*	.038	.032	−.023	−.022	.094*	.076*	−.123*	−.123*
ABILITY	−.026	−.042	−.042	.441*	.445*	−.130*	−.130*	.193*	.205*	−.118*	−.118*
PARATT	−.295			−.092*	−.092*	−.017	−.017	.292*	.292*	.015	.015
GRADETM	−.214										
DATEIND	.372										
SCHLATT	−.367										
VOCORNT	.042										
INVOLVE	−.248										
COMMIT	.083										
PEERATT	.000										
BELIEF	−.319										

for school attachment is much larger than the coefficient for the path from parental attachment to belief. This result is explained in part by an examination of Table 8, in which the associations are decomposed. Table 8 shows that parental attachment affects school attainment, which is strongly related to the level of belief, implying that part of the effects of parental attachment on belief is transmitted through school attachment.

DISCUSSION

Several limitations of the present research require comment. First, we assessed the contributions of elements of the bond in a way which parallels as closely as possible Hirschi's (1969) original research. The self-report delinquency measure used in Hirschi's research included items like "Have you ever taken a car for a ride without the owner's permission?" Such an item taps delinquent behavior for a time period prior to the collection of data. We have followed suit here and used retrospective self-reports of delinquent behavior collected with measures of elements of the bond. Our measure of delinquent behavior asks for reports on behavior over the past three years, but the school performance measure (grades) refers only to the past year. This means that the causal ordering implied by our path models is, for these data at least, questionable. This is a limitation which undermines confidence in causal interpretations in the present research. In the future, longitudinal analysis should be directed at making explicit the causal ordering of variables.

A second potential limitation is the use of a single, global measure of delinquency. This measure contains some items pertaining to delinquent behavior in school, and it is possible that the apparent influence of attachment to school and other school-linked variables on delinquency may be due only to the inclusion of these items. At the same time, however, Gold (1970) and Faine (1974) have carefully examined the dimensionality of self-report data similar to the kind used here, and have concluded that little is to be gained by use of more than a single dimension.

Despite these limitations, the development of the revised model produced several important results. The revised model implies that the low correlations between social class and ability and delinquency and the emergence of the positive direct effects reported are spurious in the context of other variables which are included in the equation to explicate the relationship to delinquency. Therefore, a substantive interpretation of the effects of social class and ability will not be offered. Parental attachment and school attachment have a strong negative relation with delinquency, whereas for grades the coefficient was moderately negative, as would be predicted by control theory. Dating was strongly related to delinquency, indicating that those boys who dated more were involved to a greater extent with delinquency.

The pattern of results which emerged for our revised measures of commitment and involvement is important. Our commitment variable did not exhibit the strong negative effects predicted by Hirschi's control theory; thus, although the present results confirm a negative association of adherence to conventional "success" or

INVOLVE		COMMIT		PEERATT		BELIEF		DELINQ		
Total	Dir.	Total	Dir.	Total	Dir.	Total	Dir.	Total	Dir.	Alpha
.094*	.081*	−.033	−.028	−.010	−.014	.009	−.039	.044	.095*	
.013	.022	.051	.047	.046	.049	.217*	.127*	−.044	.096*	.78
−.222*	.222*	−.090*	−.090*	.066*	.066*	.229*	.090*	−.298*	−.182*	.76
						−.025	−.025	−.097*	−.100*	
						−.091*	−.091*	.334*	.323*	.92
						.500*	.500*	−.239*	−.178*	.85
						.033	.033	.011	.015	.54
						−.022	−.022	−.082*	−.085*	.77
						.011	.011	.033	.034	.59
						.001	−.001	.004	.004	.62
								−.125*	−.125*	.87

* $p \leq .01$.
** $p \leq .05$.

attainment goals with delinquency, that reduced relationship may be considered unimportant. This may be due to the redundancy of commitment with other measures in the expanded models or to its low reliability. The strength of the involvement relationship is also reduced, but still significant and in the predicted direction.

The moderate and significant negative path coefficient for belief in our complex model implies that when other variables are considered simultaneously, conventional value orientations are negatively related to the incidence of delinquent behavior. In short, a lack of conventional value orientations is important in the explanation of delinquency.

Our model subscribes to the validity of the component concepts introduced in *Causes of Delinquency,* but questions the utility of that particular set of elements of socialization. In the context of statistical controls for ability, social class, and grades in school, the bond elements which emerge as important explanatory variables are attachment to parents, dating, attachment to school, belief, and involvement. A model incorporating these bond elements appears more isomorphic with theories of adolescent socialization which treat education as important in the integration of the youth into adult social life.

In considering how all the elements of the bond operate simultaneously, a different picture emerges than when applying simpler forms of analysis. Our examination of the total association, or the zero-order correlations, shows that large correlations with Hirschi's four bond elements do exist (with the exception of the element of commitment to college and a high-status career). When those components in the complex model are considered simultaneously and with controls for ability and school grades, however, it can be seen that several components are more important than others. Our results imply that models such as those depicted in Figures 2 and 3 will be more adequate and parsimonious than that originally formulated by Hirschi.

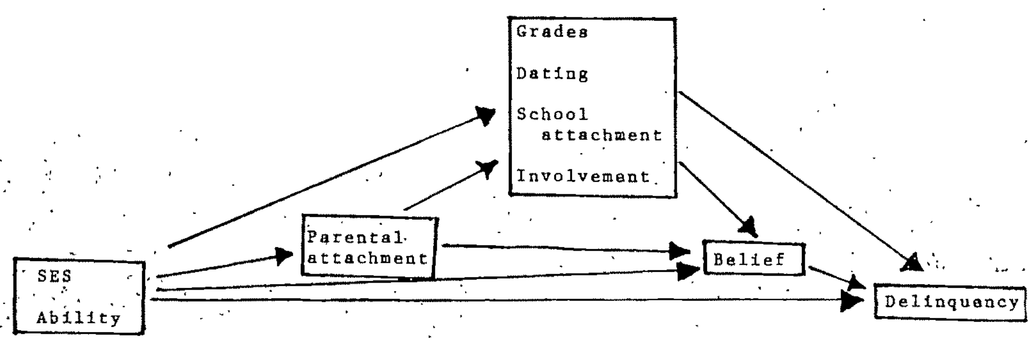


Figure 3. Modified Complex Model of the Social Bond

Table 9. Standardized and Unstandardized Path Coefficients for Revised Model of Social Control Theory

	PARATT	GRADETM1	DATEIND	SCHLATT	VOCORNT	INVOLVE	COMMIT	PEERATT	BELIEF	DELINQ
SES	.062	.032	−.022	.076	−.123	.081*	−.028	−.014	−.039	.095
	(.001)*	(.003)	(−.027)	(.002)*	(−.003)*	(.002)	(−.000)	(−.000)	(−.000)	(.003)*
ABILITY	−.042	.445	−.130	.205	−.118	.022	.047	.049	.127	.096
	(−.038)	(1.680)*	(−7.142)*	(.270)*	(−.104)*	(.024)	(.031)	(.046)	(.070)*	(.015)*
PARATT	—	.092	−.017	.292	.015	.223	−.090	.066	.090	−.182
	—	(.378)*	(−1.000)	(.415)*	(.013)	(.268)*	(−.063)*	(.067)*	(.054)*	(−.031)*
GRADETM									−.026	−.100
									(−.004)	(−.004)*
DATEIND									−.091	.323
									(−.001)*	(.001)*
SCHLATT									.500	−.178
									(.209)*	(−.021)*
VOCORNT									.033	.015
									(.020)	(.027)
INVOLVE									−.022	−.085
									(−.011)	(−.012)*
COMMIT									.011	.034
									(.009)	(.008)
PEERATT									.014	.004
									(.001)	(.001)
BELIEF									—	−.125
									—	(−.036)*
R^2	.003	.218	.020	.146	.041	.059	.010	.006	.318	.325
Residual	.998	.782	.989	.924	.975	.970	.995	.997	.826	.822

* $p \leq .01$.
** $p \leq .05$.

The emergence of schools as primary socializing institutions reflects the division of labor in a complex society, where specialized social institutions perform functions of education, socialization, and preparation for adult social roles which previously occurred within the family (Parsons, 1959; Smelser and Halpern, 1978). The bonds to society are thus formed not only in the family, as Hirschi argues, but in part in an educational context, as this research demonstrates. Social class and ability are included in our final model because of their relationship with elements of the bond which in turn affect delinquency involvement: social class and ability are treated as exogenous variables which affect both parental attachment and a set of school-related components of the social bond, which in turn affect belief. Belief was placed after the parental and school-related items because of its theoretical relationship to familial socialization and empirical relationship to the educational measures discussed earlier. In turn, the background, parental, school, and belief measures predict delinquency involvement. The resulting model of the social bond is more complex than that shown in Figure 1 and accords with theory and research on the effects of schooling and adolescence.

APPENDIX A

LIST OF ABBREVIATIONS AND SAMPLE ITEMS

The following items were used in the study. Sample items which were used to constitute scales are presented below.

1. CLOSMOM = Index to Closeness to Mother
   a. How close do you feel to your mother?
   b. How much do you want to be like the kind of person your mother (or female guardian) is?
2. CLOSFATH = Index of Closeness to Father
   a. How close do you feel to your father?
   b. How much do you want to be like your father (or male guardian) when you are an adult?
3. HOWIMPF = How important are Friends
   a. How important would you say friends are in your life?
4. TIMWFRN = Time with Friends
   a. How important is it to spend time with your friends?
5. POSSCHI = Positive School Attitudes Index
   a. I feel satisfied with school because I learn things I want to know.
   b. I believe school will help me be a mature adult.
6. NEGSCHI = Negative School Attitudes Index
   a. School is very boring for me, and I'm not learning what I feel is important.
   b. I feel the things I do at school waste my time more than the things I do outside of school.
7. ACAACHI = Academic Achievement Index
   a. Studying constantly in order to become a well educated person.
   b. Studying hard to get good grades in school.
8. ABILCON = Self-concept of Academic Ability
   a. How do you rate yourself in school ability compared with those in your grade at school?
   b. How close do you come to doing the best work you are able to do in school?
9. TCHINTER = Teacher Interest
   a. How often do teachers take an interest in my work?
10. DUNASPI = Duncan Ranking of Aspired Occupation
11. CLAROCPL = Clarity of Occupational Plans
    a. How likely is it that your plans will work out this way?
12. RCVOJT = Receive On-The-Job-Training
    a. How likely are you to receive on-the-job training?
13. COMPHS = Complete High School
    a. How likely are you to complete high school?
14. RCVMILT = Receive Military Training
    a. How likely are you to receive job training in the military?
15. RCVVOC= Receive Vocational Training
    a. How likely are you to attend a technical or vocational school?
16. ATNDCOL = Attend College
    a. How likely are you to attend college?
17. MADECOLP = Made College Plans
    a. Have you made plans to attend college?
18. DATEIND = Dating Index
    a. On the average, how many evenings a week during the school year do you usually go out for fun and recreation?
    b. On the average, how often do you go out on dates?
19. TIMEHW = Time Spent on Homework
    a. About how many hours do you spend in an average week on all your homework, including both in and out of school?
20. DSCHWFR = Discuss Homework with Friends
    a. Outside of homework how often do you have discussions with friends about ideas that come up in your courses?
21. XTRASCH = Extra School Work
    a. How often are you interested enough to do more reading or other work than the course required?
22. HONESTI = Honesty Index
    a. Never cheating or having anything to do with cheating situations even for a friend.
    b. Helping a close friend get by in a tight situation even though you may have to stretch the truth a bit to do it.
23. GUILTIN = Guilt Index
    a. I do things I feel guilty about afterwards.
    b. When I do wrong my conscience punishes me.

APPENDIX A *(Continued)*

24. SES = Socioeconomic Status—equally weighted composite comprised of
    a. Father's occupational status.
    b. Parent's education.
    c. Possessions in home.
    d. Number of books in home.
    e. Number of persons per room.
25. ABILITY = Ability Index
    a. GATB Vocabulary level.
    b. GATB Arithmetic level.
26. GRADETM = Grades Time 1
    a. What was the average grade you got in your class last year?
27. Delinquency Index Items
    a. Stayed out later than your parents said you should.
    b. Taken something not belonging to you worth under $50.
    c. Hurt someone badly enough to need bandages or a doctor.
    d. Taken an expensive part of a car without permission of the owner.
    e. Taken part in a fight where a bunch of your friends are against another bunch.
    f. Taken something not belonging to you worth over $50.
    g. Used a knife or gun or some other thing like a club to get something from a person.

APPENDIX B

ITEMS USED TO CONSTITUTE SCALES IN THE PATH ANALYSIS OF THE SIMPLE AND COMPLEX MODELS OF SOCIAL BOND

The following indices used in Tables 6, 7, 8 and 9 were constructed by equally weighting the composite elements derived from Appendix A.

SES	= Socioeconomic Status
ABILITY	= GATB Math and Verbal Test Scores
ATTACHMENT	= CLOSMOM + CLOSFATH + HOWIMPH + POSSCHI + NEGSCHI + ACAACHI + ABILCON + TCHINTR
COMMITMENT	= DUNASPI + ATNDCOL
INVOLVEMENT	= TIMEHW + DSCHWFR + X-TRASCH
BELIEF	= Honesty Index

The following indices used in Tables 10 and 11 were constructed by equally weighting the composite elements derived from Appendix A.

SES	= Socioeconomic Status
ABILITY	= GATB Math and Verbal Test Scores
PARATT	= Parental Attachment a. CLOSMOM + CLOSFATH
GRADETM	= Grades Time 1
DATEIND	= Dating Index
SCHLATT	= School Attachment a. POSSCHI + NEGSCHI + ACAACHI
VOCORNT	= Vocational Orientational a. RCVOJT + RCVMILT + RCVVOC
INVOLVE	= Involvement a. TIMEHW + DSCHWFR + XTRASCH
COMMIT	= Commitment a. DUNASPI + ATNDCOL
PEERATT	= Peer Attachment a. HOWIMPF + TIMWFRN
BELIEF	= Honesty Index

NOTE: The items used to construct scales were derived from those measures of the bond presented in Tables 2 and 3. Items were combined on the basis of theoretical parsimony and their contribution to the stability of the scale. Items which did not add to the scale stability and explanatory power of the scale were deleted.

REFERENCES

Alwin, Duane F. and Robert M. Hauser
1975 "The decomposition of effects in path analysis." American Sociological Review: 40:37–47.

Bachman, Jerald
1975 Youth in Transition. Ann Arbor: Inter-University Consortium for Political and Social Research.

Bachman, J. G., P. M. O'Malley, and J. Johnston
1978 Adolescence to Adulthood. Ann Arbor: Institute for Social Research

Bealer, R. E., F. K. Willits, and Peter R. Maida
1965 "The myth of a rebellious adolescent subculture." In L. C. Burchinal (ed.), Rural Youth in Crisis. Washington, D.C.: U. S. Department of Health, Education and Welfare.

Blau, Peter M. and Otis Dudley Duncan
1967 The American Occupational Structure. New York: Wiley.

Coleman, James S.
1961 The Adolescent Society. New York: Free Press.

Dentler, Robert A. and Lawrence J. Monroe
1961 "Social correlates of early adolescent theft." American Sociological Review 26:733–43.

Elliot, Delbert S. and Suzanne Ageton
1979 "Reconciling race and class differences in self-reported and official estimates of delinquency." Behavioral Research Institute: Mimeo.

Empey, LaMar T.
1978 American Delinquency. Homewood, Illinois: Dorsey.

Faine, J.
1974 "A multidimensional approach to understanding varieties of delinquent behavior." Ph.D. dissertation, University of Iowa. (University Microfilm No. 74-21, 888).

Galvin, James
1975 "Youth culture and adult success." Ph.D. dissertation, University of Oregon.

Gibbons, Don C.
1976 Delinquent Behavior. Englewood Cliffs, N.J.: Prentice-Hall.

Gold, Martin
1966 "Undetected delinquent behavior." Journal of Research in Crime and Delinquency 3:27–46.
1970 Delinquent Behavior in the American City. Belmont, Calif.: Brooks/Cole.

Gottfredson, Gary D.
1978 "Models and middles in human ecology: an examination of high school crime rates." Center for Social Organization of Schools, The Johns Hopkins University. Report No. 255.

Greenberg, David F.
1977 "Delinquency and the age structure of society." Contemporary Society 1:189–223.

Haller, William and Alejandro Portes
1973 "Status attainment processes." Sociology of Education 46:51–91.

Hansell, Stephen and Michael D. Wiatrowski
1980 "Competing conceptions of delinquent peer relations." Center for the Study of the Social Organization of Schools. Center Report 297. The Johns Hopkins University.

Hardt, Robert and Sandra Peterson-Hardt
1977 "On determining the quality of the delinquency self-report method." Journal of Research in Crime and Delinquency 14:247–61.

Hartshorne, Hugh and Mark A. May
1928 Studies in Deceit. New York: Arno Press.

Hindelang, Michael J., Travis Hirschi, and Joseph Weis
1979 "Correlates of delinquency: the illusion of discrepancy between self-report and official measures." American Sociological Review 44:995–1014.

Hirschi, Travis
1969 Causes of Delinquency. Berkeley, Calif.: University of California Press.

Hirschi, Travis and Michael Hindelang
1977 "Intelligence and delinquency: a revisionist review." American Sociological Review 42:571–86.

Nie, Nieman H., C. Hedlai Hull, Jean G. Jenkins, Karin Steinbrenner, and Dale H. Dent
1975 Statistical Package for the Social Sciences. New York: McGraw-Hill.

Nye, F. Ivan and James F. Short, Jr.
1957 "Scaling delinquent behavior." American Sociological Review 22:326–31.

Parsons, Talcott
1959 "The school class as a social system: some of its functions in American society." Harvard Educational Review 29:297–318.

Polk, Kenneth
1975 "Schools and the delinquency experience." Criminal Justice and Behavior 2:315–38.

Polk, Kenneth and Walter G. Schafer
1972 Schools and Delinquency. Englewood Cliffs, N.J.: Prentice-Hall.

Sewell, William H., Archibald O. Haller, and Alejandro Portes
1969 "The educational and early occupational attainment process." American Sociological Review 34:82–92.

Silberberg, N. E. and M. C. Silberberg
1971 "School achievement and delinquency." Review of Educational Research: 41:17–33.

Simms, Ronald L.
1978 "Comment on Hirschi and Hindelang. The meaning of the IQ-delinquency relationship." American Sociological Review 43:268–70.

Smelser, Neil J. and Sydney Halpern
1978 "The historical triangulation of family, economy, and education." American Journal of Sociology (Supplement) 84:288–315.

BLACK-WHITE DIFFERENCES IN THE EDUCATIONAL ATTAINMENT PROCESS: WHAT HAVE WE LEARNED?*

DENISE C. GOTTFREDSON
The Johns Hopkins University

This paper examines the credibility of inferences based on cross-group comparisons of regression coefficients using black-white differences in the educational-attainment process as an example. Reanalyses of data from previous studies fail to produce consistent results even when major differences in samples, models, and methods are held constant. Differences in substantive interpretations implied by analyses in which different measurement models are specified and in which other methodological sources of regression-slope fluctuation are evident are also illustrated. Regression-slope differences across groups in models of attainment provide ambiguous evidence on which to base statements about differences in the attainment process.

Five of the twenty-nine articles published in the American Sociological Review in 1980 used cross-group comparisons of regression coefficients as evidence for differences between groups in some social process, usually status attainment. Using the study of race differences in educational attainment as an example, this report shows that evidence for such differences is inconsistent across studies, that differences in the regression coefficients are subject to many artifactual sources of fluctuation, and that statistical inferences based on such differences are weak at best.

Only comparative studies of the educational attainment of black and white males were sufficiently abundant to allow for a sensible cross-study comparison. Duncan's (1968) analysis of the Occupational Change in a Generation (OCG) data provided a landmark for subsequent analyses of racial inequality in the status-attainment process. It suggested that the problem of racial inequality is twofold: blacks enter the occupational structure with an initial disadvantage (i.e., the mean level of socioeconomic status of parents is lower for blacks than for whites) and blacks do not get as high a return for their resources (i.e., the regression slope of attainment on background factors is not as steep for blacks as for whites). Duncan implicitly used black-white comparisons of regression coefficients and intercepts as evidence that even if blacks were to enter the labor force with the same "advantage" as whites they would nevertheless end up in lower prestige occupations and with lower earnings as a result of "occupational discrimination" (p. 108).

Attempts to specify in what ways the educational process differs for blacks and whites have elaborated Duncan's model to include measures of noncognitive socialization variables such as self-esteem and conformity (Porter, 1974; Portes and Wilson, 1976) as well as allocation variables such as curriculum placement (Thomas, 1980). Interpretations of race differences in regression coefficients for the models have ranged from socialization to allocation explanations (Kerckhoff, 1976). Socialization explanations assume that certain noncognitive skills or access

---

* Direct all communications to Denise C. Gottfredson, Center for Social Organization of Schools, The Johns Hopkins University, Baltimore, MD 21218.

I gratefully acknowledge the helpful comments given by Jerald G. Bachman, Jomills H. Braddock, Joyce Epstein, James Fennessey, Robert Gordon, Gary Gottfredson, Nancy Karweit, Alejandro Portes, Gail Thomas, Kenneth Wilson, and anonymous reviewers on early drafts of this paper. The research was supported in part by funds from the National Institute of Education through grant numbers NIE-G-78-0210 and NIE-G-80-0113. The data were made available by the Inter-University Consortium for Political and Social Research, and were originally collected by J. G. Bachman. The original data source, the Consortium and the Institute bear no responsibility for the analyses or interpretation in this paper.

to information promote future attainments, and that individuals or groups whose socialization has been deficient in these skills lack these personal characteristics that would enable them to work the system—that is, to translate certain resources into rewards. Allocation explanations, on the other hand, assume that individuals are assigned to social statuses partly on the basis of race and that attainment depends not on earned merit but on membership in an elite status.

Although these elaborations of the Blau-Duncan model clarified some issues about the interactions of race with other variables in the educational attainment model, they also raised new questions. Duncan's observation that the regression coefficients for blacks are in general lower than for whites has been replicated, but the specific differences between coefficients have appeared only inconsistently across studies.

The voluminous literature on race differences in regression slopes lacks integration. Because of their diversity it is inappropriate to directly compare the results of the studies. Differences across studies may result from cohort differences or from differing dates of data collection (base years range from 1960 to 1977). Other major differences include age (grade of subject at first contact ranges from 8th to 12th) and geographic region (most samples are not nationally representative).

Table 1 summarizes sample characteristics and the authors' interpretation of which coefficients differed for blacks and whites. A summary of the authors' interpretations of these differences for a few of the studies should suffice to familiarize the reader with common interpretations of the differences in regression slopes. See Gottfredson (1980) for a more detailed review of these studies.

Porter (1974) interpreted racial differences in magnitudes of coefficients across groups (see Table 1) and relative weights within groups according to Turner's (1960) distinction between "contest" and "sponsored" mobility. Porter claimed support for the notion that mobility, for blacks, depends upon sponsorship by the elite rather than performance, on the basis of his findings that for blacks, conformity to middle-class norms, but not ambition, depends on encouragement by significant others. Also, blacks' educational attainment is more dependent on conformity than it is on ambition. Conversely, ambition as well as conformity to middle-class norms are dependent on encouragement by significant others for whites, and their attainment is more dependent upon ambition than it is upon conformity. Porter interpreted the absence of a direct effect of intelligence and significant others' influence on grades and of grades on educational attainment for blacks as suggesting that the sponsored mobility of blacks depends on being chosen, not upon performance.

Portes and Wilson (1976) concluded that the earlier variables in the model—socioeconomic status, mental ability, and academic performance—were stronger predictors of attainment for whites than for blacks. But the mediating variables—self-esteem and educational aspirations—were more important or equally important for blacks. The authors interpreted these findings to mean that "for the (white) majority academic grades, apart from psychological effects, appear to 'carry along' individuals toward predictable levels of achievement. Black grades, especially those from all-black schools, appear to be more irrelevant as marks of achievement within the schools themselves and as criteria of selection for higher education" (p. 429). Later (p. 430) they conclude that while blacks move upward primarily through individual self-reliance and ambition, whites can rely on institutionalized mechanisms of support.

Kerckhoff and Campbell (1977b) attributed the weakness of the status-attainment model for blacks to the limited effect of socioeconomic status on educational outcomes, to inconsistency of academic performance for blacks, and to the greater importance for blacks of nonacademic performance (staying out of trouble in school) relative to academic performance. They interpreted these findings as being consistent with Porter's (1974) notions about sponsored and contest mobility systems. Teachers promote or sponsor those blacks who have few disciplinary problems and not necessarily

Table 1. Race Differences in Regression Coefficients of Educational Attainment Model Reported in Seven Studies—Males Only

Author (Date)	Sample	Sample size[a]		Grade at base year	Years data collected	Predictor variables included[b]	Differences reported for equation predicting:					
		Whites	Blacks				Educational attainment[c]	Level of occupat. aspiration	Level of educat. aspiration	Parental encouragement	Signif. others' influ.	Grade point avg.
Porter (1974)	Proj. Talent: (the 40.6% that responded to the follow-up questionnaire)	14891	435	12	1960–1965	GPA (self-report)	W	—			—	—
						LOA	0	—			—	0
						Conformity	0	B			—	0
						SOI	W	W			—	0
						Creativity	0	0			0	0
						APT	0	0			0	W
						FAOCC	W	0			W	0
Hout and Morgan (1975)	All 12th grade public school students in Louisville, KY	389	190	12	1973	LOA (friend)		0	—	—		—
						LEA (friend)		—	W	—		—
						PARENC		B	B	—		W
						GPA (school records)		B	B	0		—
						APT		W	W	—		0
						FAOCC		B	0	W		0
						HED		0	0	0		0
						SIBS		—	0	W		—
Portes and Wilson (1976)	Youth in Transition	1912	256	10	1966–1970	LEA	0		—		—	—
						Self-esteem	B		0		—	—
						SOI	0		0		—	—
						GPA (self-report)	W		0		0	—
						APT	0		W		W	W
						SES	0		W		W	0
Kerckhoff and Campbell (1977a)	All 12th grade community school system students in Fort Wayne, IN	987	74	12	1969	LEA			—			
						PARED			W			
						MOED			—			
						APT			W			
						GPA (school records)			0			
						Fatalism			B			
Kerckhoff and Campbell (1977b)	All 9th grade students in 5 of the 13 Fort Wayne, IN public junior high schools	390	113	9	1969–1974	LEA	W		—			—
						GPA (school records)[d]	W		W			—
						Disciplinary problems	B		0			—
						FAOCC	W		W			0
						FAED	W		W			W
						MOED	0		0			0
						SIBS	0		0			W
						APT	0		W			0

Table 1. Continued

Author (Date)	Sample	Sample size[a] Whites	Sample size[a] Blacks	Grade at base year	Years data collected	Predictor Variables included[b]	Differences reported for equation predicting: Educational attainment[c]	Level of occupat. aspiration	Level of educat. aspiration	Parental encouragement	Signif. others' influ.	Grade point avg.
DeBord, Griffin, and Clark (1977)	Those students enrolled in 23 Mississippi schools who answered all questionnaire items (46%)	1014	439	8–12	1972	LOA		—	—			—
						LEA		—	—	—		—
						TCHENC		0	0	—		—
						PARENC		W	W	—		—
						PEERPLAN		0	0	—		—
						GPA (school records)		0	0	0		—
						APT		0	B	0		0
						FAED		0	0	0		W
						MOED		0	0	B		0
						HOCC		W	0	0		0
Thomas (1980)	Subsample of National Longitudinal Sample for which standardized test data and SES data exist (64.5% of entire sample.)	4758	346	12	1972–1974	LEA	0	Information for these outcomes not reported by author				
						Perceived college ability	0					
						SOI	0					
						Curriculum enrollment	W					
						Rank in class (school records)	B					
						APT	B					
						SES	B					

NOTE: Equations for variables included in only one study are not shown on this table. No race interactions were reported in these equations except for the following: self-esteem depends on aptitude for blacks and not whites (Portes and Wilson, 1976); teacher encouragement depends more upon mother's education for blacks than for whites (DeBord, Griffin, and Clark, 1977); counselor reports of disciplinary problems depend upon measures of SES for whites and not blacks (Kerckhoff and Campbell, 1977b).

Notation in columns is as follows: 0 = no race difference reported
W = race difference favoring whites reported
B = race difference favoring blacks reported
— = model does not include this test
blank = criterion not examined

[a] For the Portes and Wilson (1976), Kerckhoff and Campbell (1977b) and Thomas (1980) studies, the sample sizes given are for the base years of the surveys used. Numbers of cases on which the correlations were based were provided only by Kerckhoff and Campbell. For that study, actual Ns for whites ranged from 296–390 and for blacks from 64–113.

[b] The following abbreviations are used in this table:

- LOA —Level of occupational aspiration (or expectation), measured in Duncan SEI prestige score
- LEA —Level of educational aspiration (or expectation)
- SOI —Significant others' influence (usually a composite based on parents', teachers' and friends' encouragement to respondent to continue schooling)
- PARENC —Parental encouragement to continue schooling
- TCHENC —Teacher and counselor encouragement to continue schooling
- GPA —Grade point average
- APT —Aptitude test score
- FAED —Father's educational level
- MOED —Mother's educational level
- PARED —Parents' educational level
- FAOCC —Father's occupational prestige (Duncan SEI score)
- HOCC —Head of household's occupational prestige
- HED —Head of household's educational level

[c] Measured by years of education attained in all studies but Thomas (1980). There it was measured by a dummy coding of attended college/did not attend college.

[d] A measure of Senior High GPA was also included in this study, but to preserve comparability only Junior High GPA is reported here.

those who perform well academically. These conclusions are based in large part on the following findings: A model including a measure of disciplinary problems and two measures of grade-point average (one from junior and one from senior high school) predicted educational attainment better for blacks than for whites. Measures of status-background variables had no significant direct effects on any outcome for blacks with the exception of a direct effect of mother's education on early grade-point average. Grades in senior high school were more predictable from grades in junior high school for whites than blacks, and a measure of disciplinary problems was more important than a measure of grade-point average for explaining blacks' attainment, while the reverse order of effects was observed in the white equation.

No cross-study agreement exists about specific race differences in the parameters of this model of educational attainment. Major discrepancies beyond those which are obvious from Table 1 exist among the studies. Whereas Portes and Wilson (1976) found educational expectations to be a strong predictor of attainment for blacks, Kerckhoff and Campbell (1977b) found no effect of aspirations on attainment for blacks. While both Portes and Wilson (1976) and Porter (1974) found school performance to be insignificant for explaining attainment for blacks, Kerckhoff and Campbell (1977b) found that senior-high-school GPA was the only variable whose coefficient reached the $p < .1$ level of significance for blacks. Other discrepancies involve the determinants of educational expectations. Some studies (Hout and Morgan, 1975; Kerckhoff and Campbell, 1977b) implied that academic performance, but not mental ability, was predictive of educational expectations for black males. In contrast, DeBord, Griffin, and Clark (1977) found school performance to be significant for all groups except black males and the effect of ability to be greatest for black males. Similar discrepancies exist with regard to the influence of significant others. Studies that combined the sources of influence found no effect for black males on their educational expectations. Hout and Morgan found peer effects for all groups but black males and parental encouragement effects for all groups but strongest for black males. (They also interpreted the large GPA effect as an indication of a strong teacher-encouragement effect for black males.) DeBord, Griffin, and Clark found almost the opposite. All three influences were significant for blacks, but encouragement from parents was much more predictive of educational expectations for whites than for blacks. Finally, the only difference that appeared with any consistency across studies—that academic performance is more important for predicting the educational attainment of whites than blacks—was disconfirmed by Thomas' study.

RESOLVING DIFFERENCES AMONG STUDIES

As Jencks et al. (1979) recently demonstrated, differences in results based on different surveys may result from a myriad of seemingly arbitrary decisions that must be made by both individual researchers and survey organizations. Differences in sampling frame, measurement procedures, attrition, and categorization of data cause means, standard deviations, and associations among the same variables to differ from survey to survey. At the level of the individual researcher, choice of population, definition of variables, recoding and transformation of variables, and treatment of missing data are among the decisions that affect research results. Researchers are seldom aware of the ways in which "seemingly innocuous 'procedural' or 'methodological' decisions affect outcomes" (p. 289). Jencks et al. conclude that "surveys agree well on the broad, general picture, but detailed interpretations must still be treated with some caution" (p. 282).

Several major sampling and procedural differences among the studies reviewed are likely causes of their discrepant results. First, the use of different, often incorrect, criteria for identifying cross-race differences in regression coefficients is a likely cause of failure to replicate, as is the use of statistics which assume random observations with nonrandom samples. Second, the use of small subsamples of blacks

results in unstable regression coefficients and increases the likelihood of finding cross-group differences when the sampling variation of coefficients is not appropriately considered. Differences in the constructs included in the causal models on which the equations are based and in the particular operational measures used for the constructs also affect results. Finally, differences across groups in measurement reliabilities present a potential threat to the interpretability of group differences in regression coefficients. Each of these sampling characteristics and procedural decisions affect the values of regression coefficients and interject ambiguity into substantive interpretations based on the regression-slope differences.

Criteria for Identifying Cross-Group Differences

A difference interpreted as substantively meaningful in one study may be regarded as noise in another. The variety of criteria used in the studies reviewed for identifying regression-coefficient differences implies that the choice of an appropriate statistic is not always obvious. At least six different criteria were used in the seven studies reviewed here: One study required that the difference between the black and white coefficients exceed one standard error of the white coefficient; another required that it exceed 1.5 standard errors of the black coefficient. One study used a t-test to determine whether the difference could have arisen as a result of random fluctuation, given the white coefficients as population values. One used an F-test to see whether adding an interaction term for race times a given variable added significantly to the variance explained in the criterion. One study chose an arbitrary value and required the difference between the standardized regression coefficients for blacks and whites to exceed it before being considered important, and another simply used the "eyeball" method for detecting differences.

Standard statistical texts discuss appropriate techniques for identifying regression-coefficient differences. Kerlinger and Pedhazur (1973, Chapter 8) and Hanushek and Jackson (1977, Chapter 4) suggest the use of interaction terms computed by multiplying the grouping variable (in this case race) by the predictor of interest. The statistical significance of the interaction term can be assessed with a t-statistic to test the null hypothesis that the interaction term's coefficient (if only one term is being tested) equals zero, or with an F-statistic to test the null hypothesis that adding the interaction term to the equation adds nothing to the explained variance of the criterion. Another appropriate statistical test of the null hypothesis of equal slopes is a t-test for random variables with unequal variances.[1]

In all but one of the studies reviewed, the criteria used for identifying differences in coefficients were less stringent than is appropriate, resulting in the rejection of the null hypothesis of no difference too often.

Departures from randomness in the sampling designs for surveys add another source of nuisance fluctuation to regression coefficients when the sampling characteristics are not taken into account in the analysis stage. In particular, nonindependence of observations (implying unequal or correlated error variance across observations) biases the results of standard statistical tests by causing standard errors to be underestimated and, again, too often rejecting the null hypothesis of no difference. The sampling designs for both the Youth in Transition and the National Longitudinal Surveys, two large-scale, national samples used in the studies reviewed, were multistage cluster designs resulting in nonindependence of observations.

Table 2 demonstrates the consequences of violating the nonindependence as-

---

[1] This test statistic is

$$t = (b_b - b_w)/(\text{st. err}_b^2 + \text{st. err}_w^2)^{1/2}$$

A correction for the degrees of freedom reflecting unequal sample sizes (Hays, 1966) can also be used. Other tests are possible. For example, a goodness of fit of the overall model to the data might be used as the criterion for accepting or rejecting the notion of statistical interaction. The fit for models specifying equality constraints on parameters across groups can be compared to that for models which allow parameters to be estimated separately for each group. Such a test is possible with LISREL IV (Joreskog and Sorbom, 1978).

Table 2. Proportions of Variance Explained and Increments to Explained Variance Due to Race Interactions and Random Grouping Variable Interactions

		Criterion				
		Academic performance	Significant others' influence	Level of educational aspirations	Level of occupational aspirations	Educational attainment
R^2 for model with no interactions		.199	.133	.198	.279	.481
R^2 for model with all race interactions		.222	.140	.206	.284	.490
Increment due to race interactions		.023**	.007*	.008**	.005	.009**
Increment due to random grouping variable interaction for group number:	1	.000	.013**	.014**	.018**	.007**
	2	.001	.009**	.003	.003	.007**
	3	.002	.006*	.002	.003	.012**
	4	.001	.004	.003	.003	.005*
	5	.000	.000	.001	.005	.056**
	6	.001	.003	.005*	.013**	.001
	7	.000	.000	.003	.002	.002

* $p<.05$.
** $p<.01$.

sumption. It shows proportions of variance explained by additive models and models including race interaction terms for each equation implied by the Wisconsin model of status attainment. Data are from the Youth in Transition Project (Bachman, 1975). See Appendix, Table A, for descriptive data on the observed measures.

The significance level of increments to explained variance of the interaction model over the additive model are tested with an F-test. This test is not useful for testing hypotheses about specific interactions but alerts the researcher to the possibility of an interaction with one or more of the variables in the equation. Such tests are found frequently in comparative status-attainment studies and yield results similar in terms of magnitude of increments to those in the top panel of Table 2.

The bottom panel of Table 2 shows parallel information for increments due to interactions with random variables. All whites in the YIT sample (N = 1912) were assigned a random-grouping variable so that seven nonoverlapping groups with Ns ranging from 239 to 289 could be identified and the significance of interaction terms computed using these random-grouping variables was tested.

The table shows that increments due to the random interaction terms often are larger than those due to race interactions. Clearly, using the standard F-test is inappropriate with these data. The test should reject the null hypothesis (that the addition of the interaction terms does not increase the prediction of the criterion) in 1.75 of the 35 random tests at the .05 level and .35 times at the .01 level. Instead it is rejected 16 times at the .05 level and 9 times at the .01 level.

Cross-group comparisons of specific regression coefficients using a standard t-test (see footnote 1) yield equally questionable results. Coefficients estimated for random white subgroups diverge as much or more from the values estimated for the full white sample than do those estimated for the black subsample. Clearly, the use of standard statistical tests with these data yield misleading results.

It is sometimes possible to adjust for heterogeneity of error variance using techniques familiar to statisticians. Generalized least squares (Hanushek and Jackson, 1977, Chapter 6) is one such technique; design effects (a measure of the extent to which standard statistical formulas underestimate actual standard errors) is another. Bachman, O'Malley, and Johnston (1977, Appendix B) provide approximations for standard errors adjusted for nonindependence using a design effect that was calculated for the entire Youth in Transition sample. They chose not, however, to provide a design effect that could be used for the black subsample, where the problem of nonindependence is much more severe (over two-thirds of the blacks are located in only 8 of the 87 schools sampled).

The original investigators of the Youth

in Transition data have consistently and repeatedly discouraged the study of race differences using these data because the generalizability of results from the black subsample is severely limited. Differences within the black subsample based on differences in school and community environments are great, and the likelihood is high that observed differences between the blacks in the sample and other subgroups result from school differences rather than race differences because the blacks are clustered in only a few, primarily segregated schools.

Bachman et al.'s design effect was used to adjust the standard errors of regression coefficients for the total white, black, and random white subgroups in the exercise described above. When the t-tests are repeated after making this correction, most differences between coefficients for the total white and all other groups are reduced to nonsignificance.

The researcher's decision about how to test for the presence of group differences in regression coefficients has serious consequences for the study's conclusions, and the problem is worse when data do not abide by the assumptions required by standard statistical tests. The choice of an appropriate statistic is not always obvious, though. Had I adopted the criteria for identifying race differences in regression coefficients used in the studies reviewed earlier, I would have concluded that substantively interesting race differences exist. Instead I conclude that the observed differences are, by and large, due to characteristics of the sample that result in inefficient estimates.

Small Subsamples and Different Causal Models

In an effort to uncover cross-study consistencies by controlling for size of subsamples and differences in causal models used, I reanalyzed data from only those studies based on subsamples of at least two hundred cases using a common causal model and method for identifying regression coefficient differences.[2]

Regressions were computed separately for each racial group using the original Wisconsin model (Sewell, Haller, and Portes, 1969). This entailed excluding variables such as "conformity to middle-class norms" and "self-esteem" which are unique to one or another study. Also, for those studies that used disaggregated measures of some of the model's constructs (DeBord, Griffin, and Clark, 1977; Thomas, 1980), I included all measures as indicators of the constructs in the Wisconsin model (LISREL was used for all reanalyses). Multiple indicators were used for the Significant Others' Influence construct—reports by teachers, peers, and parents—and for the Socioeconomic Status construct—measures of parents' education, fathers' occupation, and (in Thomas only) a Household Index.[3]

Because the reanalyses are based on published correlation matrices, it is not possible to follow the procedures suggested by Kerlinger and Pedhazur (1973) or Hanushek and Jackson (1977) to test for statistically significant increments to explained variance in the criterion due to the interaction terms. Instead, I chose to test the null hypothesis that the regression coefficients for the black and white subsamples are equal (in each study) with a t-statistic for random variables with unequal variances.

This is not an entirely appropriate test. It does not explicitly correct for the nonhomogeneity of variance across observations caused by the cluster-sampling designs used in many of the studies, although the use of separate variance estimates for blacks and whites helps somewhat to offset this problem to the extent that blacks and whites are highly segre-

---

[2] Published correlation matrices, means, and standard deviations were available for all but one of the studies included in this reanalysis. The data for the study that did not publish the necessary information (Thomas, 1980) was supplied by its author.

[3] The choice of indicators for each of the constructs was determined solely by the requirement of equivalent measures across all studies. The intent of this exercise is to eliminate obvious differences among the studies in order to uncover consistencies rather than to estimate true parameters for the status-attainment model. Hence, although different specifications (such as inclusion of other psychological variables or disaggregating significant others' influence) are superior, some correctness must be sacrificed here.

gated in schools (which is the case at least in the Youth in Transition data). This test is also flawed because it assumes that each test is independent of other tests, which is not the case because the values of the regression coefficients in a given equation are not independent. Despite these problems, I considered it to be acceptable for the present task, which is to apply a uniform criterion for identifying coefficient differences to the results of several studies rather than to learn about actual race differences in regression coefficients.

Table 3 summarizes the results of the reanalyses of four studies and shows that little consistency emerges. Some of the differences reported in the original studies are not upheld in the reanalysis. All but one of the reported differences in the effects of SES are no longer significant. Four of the seven differences found are found in only one study. The effect of Aptitude on Level of Educational Aspirations was found to differ by race in two of the three studies including that test, but the direction of the "advantage" was different in each study. The two remaining differences, Aptitude on Academic Performance, appearing in three of the four reanalyses, and Academic Performance on Significant Others' Influence, appearing in two of the studies, all indicate a larger effect for whites than for blacks.

Reliability, Validity, and Interpreting Slope Differences

Another source of variation in regression slopes is imperfect measurement of constructs. Regression coefficients are biased downward to the extent that the constructs they purport to describe are imperfectly measured. Differences across groups in the construct validity of measures can render comparisons of their regression coefficients useless. For example, in determining the effect of investing in vocational preparation on later income, choosing to measure the investment by the number of years of college completed is likely to result in a larger regression coefficient for academics than for machinists. Perhaps machinists do not receive as high a rate of return on their investments as do academics, but more likely, the construct—investment in vocational preparation—is not well measured by college attendance for machinists. Thus, the regression coefficient for machinists is biased downward.

Bielby, Hauser, and Featherman (1977) investigated the consequences of ignoring differential measurement error across groups. They used reports of socioeconomic background and educational attainment collected at three times, systematically varied the specification of the measurement properties for each group, and examined the fit of the data to each model. They found more measurement unreliability among the black than the white subsample, and found that ignoring measurement error led to misinterpretations of their data.

Black and white males who participated in the base year (1972) of the National Longitudinal Study (NLS) of the high-school senior class of 1972 were used to illustrate the consequences for substantive interpretations of different specifications of the measurement model. These are the same data used in Thomas' study except that the present study included the respondents who did not have data for the standardized mental abilities tests (37% of the black and 28% of the white males) and computed pairwise present correlation coefficients. Thomas' study excluded those cases from all analyses. (Means, standard deviations, and correlations for these data appear in Appendix, Table D.)

Equality constraints across black and white subgroups were manipulated for segments of the model (measurement and structural) shown in Figure 1,[4] and the relative goodness of fit of the different models was assessed. Differences between the black and white coefficients for the variables in the equation predicting academic performance—the one equation in which black-white differences in regression coefficients were replicated—were also examined for the various specifications.

[4] This can be accomplished using LISREL IV (Joreskog and Sorbom, 1978). When an equality constraint for any parameter is imposed, the groups are analyzed simultaneously holding that parameter constant across the groups.

Table 3. Unstandardized Regression Coefficients for Wisconsin Model

Predetermined variables	Criterion variables: Academic performance		Significant others' influence		Level of educational aspirations		Educational attainment	
	Blacks	Whites	Blacks	Whites	Blacks	Whites	Blacks	Whites
Data from Porter (1974); Black N = 435, White N = 14891								
Socioeconomic status	.018	.008*	.029	.074**	(no measure of educational expectations included in this study)			
Aptitude	.006	.049**[b]	.068**	.074**				
Academic performance	—	—	.016	.282**[b]				
Significant others' influence	—	—	—	—				
R^2	.006	.113	.133	.191				
Data from DeBord, Griffin, and Clark (1977); Black N = 439, White N = 1014								
Socioeconomic status	−.011[a]	.076*	.127**	.082**	−.023	.018	(no measure of educational attainment included in this study)	
Aptitude	−.155**	.161**	.002	.013*	.126*	−.007[b]		
Academic performance	—	—	.047	.052**	.059	.181*		
Significant others' influence	—	—	—	—	3.147**	6.266**[b]		
R^2	.240	.279	.188	.435	.383	.653		
Data from Portes and Wilson (1976); Black N = 256, White N = 1957[c]								
Socioeconomic status	−.041	.022	.055*	.094**	.011	.024**	.027*	.027**
Aptitude	.499**	1.637**[b]	.041	.210**[a]	−.044	.035*[b]	.115*	.097**
Academic performance	—	—	.094*	.064**	.046**	.038**	.026	.062**[a]
Significant others' influence	—	—	—	—	.055*	.059**	.034	.040**
Level of educational aspirations	—	—	—	—	—	—	.452**	.404*
R^2	.056	.271	.066	.186	.116	.347	.279	.433
Data from Thomas (1980); Black N = 346, White N = 4758[c]								
Socioeconomic status	−.156	.029	.029*	.025**	.109*	.138**	.057*	.039**
Aptitude	.428**	.624**[b]	.015*	.008**	.128**	.092**	.019	.017**
Academic performance	—	—	.001	.010**[b]	.048*	.073**	.025*	.022**
Significant others' influence	—	—	—	—	2.091*	2.202**	.789*	.393**
Level of educational aspirations	—	—	—	—	—	—	.154**	.207**
R^2	.113	.268	.183	.168	.319	.470	.378	.474

* Coefficient is significant at the $p < .05$ level.
** Coefficient is significant at the $p < .01$ level.
[a] Difference between black and white coefficients is significant at the $p < .05$ level.
[b] Difference between black and white coefficients is significant at the $p < .01$ level.
[c] Sample sizes given are for the base years of the surveys used. Numbers of cases on which the correlations were based were not provided.

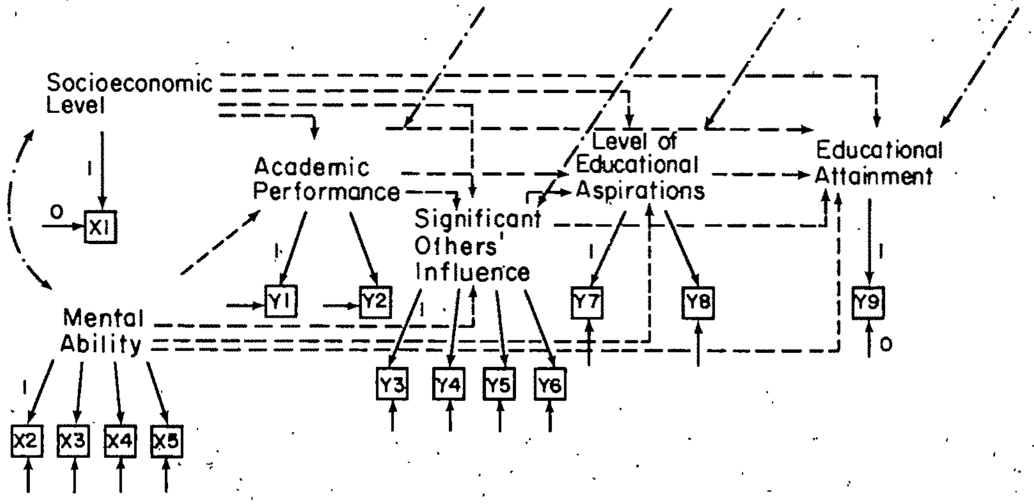


Figure 1. LISREL Model of Educational Attainment

NOTE: Descriptions of measures along with their means, standard deviations and correlations for black and white subgroups appear in Appendix. Dotted lines — — — represent "structural" paths, solid lines represent "measurement" paths, and irregular lines (—.—.—) represent unanalyzed correlations or residuals.

Model A (See Table 4) fits the data better ($\chi^2_{D\text{-}A}$ = 1450, df = 41) than Model D, which treats data from the two subsamples as though they were from a single population. This improvement in fit implies that at least some element of the model differs for the two groups. Models B and C are attempts to narrow down the possible locations of the differences. Model B constrains all structural coefficients (those marked with a solid line in Figure 1) to be equal for the black and white subsamples and allows the measurement coefficients (marked with dotted lines in Figure 1) to be estimated separately for each group. Model C reverses that procedure. Relaxing the equality constraints on the measurement model (Model B) improves the model's fit (over Model D) more (χ^2_{D-B} = 1382, df = 27) than does relaxing the equality constraints of the structural parameters (over Model D) (χ^2_{D-C} = 447, df = 21). In other words, allowing only the measurement of the constructs to vary across groups accounts for 95% of the difference between the best and worst fitting models, while allowing only the structural coefficients to vary accounts for only 31% of that difference.

Thomas' analysis of the NLS data (See panel 4 of Table 3), which specifies perfect measurement of the constructs, implied that the effect of mental ability on academic performance is greater for whites than for blacks. The regression coefficients reported in Table 4 show that once estimates are "corrected" for differential measurement reliability (Model A), the oft-observed black-white difference in the effect of mental ability on academic performance disappears. Only when the model forces all black-white differences in the correlations of the observed variables to be represented in the structural coefficients (Model C) does it reappear.

The inherent weakness of the preceding analysis should be pointed out.[5] My specification of the LISREL model assumes that the same variables measure the same constructs for blacks and whites. It represents low correlations among the indicators of a construct as measurement error. An alternative hypothesis is that the choice of indicators for a construct is inappropriate for one group or another, as in my earlier vocational preparation example. Perhaps blacks live in a less orderly world than do whites, and the low correlations among, for example, teachers', parents', and peers' influences are indicative of a real race difference in

[5] I wish to thank an anonymous reviewer for clarifying this weakness.

Table 4. Relative Goodness of Fit and Regression Coefficients for Models of Educational Attainment Allowing for Imperfect Measurement

Model name and description	χ^2 goodness of fit, corrected for degrees of freedom[a]	Unstandardized regression coefficients;[b] Academic performance on: Socioeconomic status, Black	Socioeconomic status, White	Mental ability, Black	Mental ability, White
A: No equality constraints across groups	2974	−.017 (.012)	−.014 (.004)	.232 (.018)	.252 (.005)
B: Structural coefficients constrained to be equal across groups	3042	−.015 (.004)		.251 (.005)	
C: Measurement coefficients constrained to be equal across groups	3977	−.018 (.012)	−.014 (.004)	.174[c] (.012)	.258 (.005)
D: Structural and measurement coefficients constrained to be equal across groups	4424	−.013 (.004)		.249 (.005)	

[a] This statistic is computed by subtracting the models' degrees of freedom from their χ^2 values.
[b] Standard errors are in parentheses.
[c] Difference between black and white coefficient is statistically significant at the $p < .001$ level.

homogeneity of expectations from different sources of influence. While this alternative hypothesis seems reasonable for explaining low correlations among the indicators of significant others' influence for blacks, it is less plausible as an explanation of the lower correlations for blacks than whites among academic ability and performance variables. Critical tests that distinguish among the hypotheses that predict lower correlations among indicators for blacks than whites are needed to strengthen arguments for either perspective.

SUMMARY AND CONCLUSION

Substantive inferences about race differences in the educational attainment process are unwarranted on the basis of differences in regression coefficients in the published literature. Studies examining racial differences in the attainment process over the past decade do not agree on the nature of the differences observed, and inconsistencies persist even when major differences in the samples, models, and methods are held constant. Additional doubt is cast on substantive interpretations of the observed differences when different specifications of the measurement model are shown to imply different substantive interpretations.

Conclusions of studies which purport to find group differences in structural equations parameters can be strengthened in two ways. First, rigorous model development is essential. Both measurement and structural models should be generated by specific theories, and alternative theories that predict the same outcome must be systematically ruled out. A theory that hypothesizes race differences in structural parameters is likely also to hypothesize differences in the measurement model.

Second, assumptions required by statistical tests must be taken seriously. The consequences of sampling procedures that result in nonrandom samples must be carefully considered and compensated for with appropriate statistical tests.

Evidence from studies about group differences in social processes will remain ambiguous until these measures are taken. Designing creative studies that focus better on the research questions generated by specific theories holds more promise than do continuing efforts to squeeze precise parameter estimates from data that generally lack the power to address the questions.

APPENDIX

Table A. Descriptions of Variables in Wisconsin Model of Educational Attainment: YIT Data

Abbreviation	Description
SEL	Socioeconomic Level: a summary index, consisting of six equally weighted components: father's occupational status, father's educational level, mother's educational level, number of rooms per person in home, number of books in home, checklist of other possessions in home. See Bachman (1970, Appendix B) for a detailed description of the construction and validity of this scale.
MA	Mental Ability: factor score from a principal components factor analysis of three ability tests: 1) Ammons Quick Test of General Intelligence; 2) GATB, part of J-Test of Vocabulary Level; and 3) Gates Reading Comprehension.
AP	Academic Performance: respondent's report of his average grade received in his classes for the past year.
SOI	Significant Others' Influence: an index computed on the basis of two questions: "How do these people feel about whether you should go to college?" and "What if you decided *not* to go to college—how would they feel?" A score of "3" was given if the respondent was being encouraged to attend college and bad feelings would result from nonattendance, a score of "2" was given if the respondent was being encouraged but the referrent wouldn't care if he or she decided not to attend, and a score of "1" was assigned if the respondent was not being encouraged to attend college. Scores for questions referring to father, mother, teacher, and friend were summed and the resultant composite score was standardized.
LEA	Level of Educational Aspirations: this index is based on responses to the questions regarding short-range educational plans. A code of "3" was assigned if the respondent definitely planned to attend graduate or professional school, "2" if he definitely planned to attend college, "1" if he definitely planned to complete high school, and "0" otherwise.
LOA	Level of Occupational Aspirations: Duncan prestige ranking for the occupation named by respondent in response to the question, "In the long run, what sort of work do you think you might do for a living?"
EDATT	Educational Attainment: an index constructed by adding standardized scores for the following variables. a) Information about educational pursuits completed or in progress recoded into a scale ranging from "0" (have not yet completed high school or earned a high-school equivalency), through "6" (have attended or am attending a graduate or professional school after college). b) "How many years of schooling have you completed?" c) "What is the highest degree you have earned?"

NOTE: LEA and LOA were constructed using responses from time 2, EDATT from time 5, and all other variables from time 1.

Table B. Means, Standard Deviations, and Pairwise Present Correlations for Variables in Wisconsin Model of Educational Attainment: YIT Data

	SEL	MA	AP	SOI	LEA	LOA	EDATT	$\bar{x}$	SD	Valid N
SEL	—	.435	.252	.312	.348	.350	.428	509.747	77.219	1868
MA	.467	—	.482	.313	.331	.469	.492	.154	.801	1912
AP	.058	.247	—	.264	.377	.403	.510	40.239	7.431	1904
SOI	.192	.097	.177	—	.277	.317	.308	5.680	2.654	1433
LEA	.104	.002	.171	.135	—	.427	.497	1.401	.770	1623
LOA	.248	.312	.143	.148	.265	—	.507	59.830	24.709	1318
EDATT	.301	.396	.256	.246	.353	.270	—	.202	2.891	1359
$\bar{x}$	449.973	−1.057	38.254	5.484	1.350	53.282	−1.431			
SD	78.084	1.063	6.463	2.843	.895	24.674	2.174			
Valid N	226	250	256	188	197	181	119			

NOTE: Statistics for blacks appear below the diagonal and for whites above the diagonal.

Table C. Descriptions of Observed Variables in Model of Educational Attainment (Figure 1): NLS Males

Symbol in figure	Abbreviation	Description
X1	SES	Index of socioeconomic status resulting from factor analysis of father's education, mother's education, parents' income, father's occupation, and household items.
X2	VOCAB	Vocabulary subtest score (standardized).
X3	PICTURE	Picture Number subtest score (standardized).
X4	READING	Reading subtest score (standardized).
X5	LETTER	Letter groups subtest score (standardized).
Y1	SLFGPA	Respondent's report of grades so far in high school.
Y2	GPA	Grade-point average from school records.[a]
Y3	FAEX	Response to a question asking how much schooling respondent's father wants respondent to get. Responses range on a 6-point scale from "quit high school without graduating" to "go to graduate or professional school."
Y4	MOEX	Same as above, but refers to mother.
Y5	TCHINF	Response to a question concerning influence of teachers and counselors on college attendance. Possible responses are, "discouraged me," "didn't try to influence me," and "encouraged me."
Y6	PEERPLAN	Respondent's report of his or her close friends' plans for next year. Coded "2" for college, "1" for vocational, technical, business or trade school, and "0" otherwise.
Y7	EDASP	Highest level of education respondent would like to attain, using same response scale as for Y3.
Y8	EDEX	Highest level of education respondent plans to attain, using same scale as for Y3.
Y9	ATTAIN	Highest level of education or training respondent attained by October, 1976. Responses range from finished high school through finished Ph.D. or advanced professional degree.

NOTE: Variables X1 through Y8 are taken from the base year questionnaire which was administered in 1972 when the students were seniors. Y9 was taken from the third follow-up questionnaire which was administered four years later.

[a] This score was imputed from reports of actual grades (letter or grade-point average) and rank in class.

Table D. Means, Standard Deviations, and Pairwise Present Correlations for Observed Variables in Model of Educational Attainment (Figure 1): NLS Males

	SES	VOCAB	PICTURE	READING	LETTER	SLFGPA	GPA	FAEX	MOEX	TCHINF	PEER PLAN	EDASP	EDEX	ATTAIN	$\overline{X}$	SD	Valid N
SES	—	.334	.157	.306	.251	.209	.200	.428	.398	.147	.338	.360	.424	.409	1.030	6.717	8532
VOCAB	.253	—	.253	.647	.428	.410	.450	.405	.407	.226	.338	.379	.411	.424	51.659	9.699	6205
PICTURE	.138	.189	—	.322	.397	.335	.332	.266	.260	.188	.221	.237	.273	.293	49.676	9.730	6205
READING	.220	.611	.298	—	.532	.423	.458	.420	.419	.264	.330	.402	.436	.422	51.548	9.497	6205
LETTER	.225	.402	.372	.548	—	.392	.412	.375	.362	.255	.321	.331	.370	.375	50.662	9.427	6205
SLFGPA	.086	.173	.168	.257	.212	—	.706	.394	.396	.286	.321	.405	.477	.447	5.389	1.423	7943
GPA	.082	.238	.253	.321	.314	.503	—	.386	.389	.277	.313	.384	.456	.450	7.082	3.047	6544
FAEX	.310	.269	.170	.323	.339	.170	.179	—	.910	.361	.448	.684	.748	.561	4.354	1.251	6544
MOEX	.328	.356	.187	.390	.360	.182	.216	.838	—	.357	.451	.682	.739	.547	4.417	1.222	6671
TCHINF	.044	.141	.083	.192	.189	.167	.138	.135	.236	—	.268	.320	.314	.267	2.612	.543	6411
PEERPLAN	.185	.260	.178	.221	.218	.108	.140	.232	.274	.164	—	.453	.507	.452	1.258	.933	7606
EDASP	.239	.293	.227	.356	.324	.236	.190	.513	.485	.162	.353	—	.732	.544	4.853	1.274	4572
EDEX	.280	.348	.135	.381	.327	.304	.246	.529	.562	.205	.371	.566	—	.646	4.215	1.384	4991
ATTAIN	.266	.319	.214	.346	.325	.241	.314	.372	.390	.205	.280	.409	.429	—	3.309	1.953	7658
$\overline{X}$	−5.949	42.078	44.086	42.042	40.906	4.901	5.646	4.283	4.328	2.683	.983	4.645	4.047	2.709			
SD	5.886	7.334	9.250	9.055	10.477	1.287	2.902	1.308	1.291	.545	.954	1.338	1.406	1.825			
Valid N	1336	859	859	859	859	1249	1032	757	893	901	1207	378	448	1106			

NOTE: Statistics for blacks appear below the diagonal and for whites above the diagonal.

REFERENCES

Bachman, Jerald G.
1970 Youth in Transition, Vol. 2: The Impact of Family Background and Intelligence on Tenth Grade Boys. Ann Arbor: Survey Research Center, Institute for Social Research.
1975 Youth in Transition Documentation Manual. Ann Arbor: Survey Research Center, Institute for Social Research.

Bachman, Jerald G., Patrick M. O'Malley, and Jerome Johnston
1977 Youth in Transition, Vol. 6: Adolescence to Adulthood—Change and Stability in the Lives of Young Men. Ann Arbor: University of Michigan, Institute for Social Research.

Bielby, William T., Robert M. Hauser, and David L. Featherman
1977 "Response errors of black and nonblack males in models of intergenerational transmission of socioeconomic status." American Journal of Sociology 82:1242–1288.

DeBord, Larry W., Larry J. Griffin, and Melissa Clark
1977 "Race and sex influences in the schooling processes of rural and small town youth." Sociology of Education 42:85–102.

Duncan, Otis D.
1968 "Inheritance of poverty or inheritance of race?" Pp. 85–110 in Daniel Moynihan (ed.), On Understanding Poverty: Perspectives from the Social Sciences. New York: Basic Books.

Gottfredson, Denise C.
1980 "Personality and persistence in education." Doctoral dissertation, The Department of Social Relations, The Johns Hopkins University.

Hanushek, Eric A. and John E. Jackson
1977 Statistical Methods for Social Scientists. New York: Academic Press.

Hays, William L.
1966 Statistics for Psychologists. New York: Holt, Rinehart and Winston, Inc.

Hout, Michael and William R. Morgan
1975 "Race and sex variations in the causes of the expected attainments of high school seniors." American Journal of Sociology 81:364–94.

Jencks, Christopher, Susan Bartlett, Mary Corcoran, James Crouse, David Eaglesfield, Gregory Jackson, Kent McClelland, Peter Mueser, Michael Olneck, Joseph Schwartz, Sherry Ward, and Jill Williams
1979 Who Gets Ahead? The Determinants of Economic Success in America. New York: Basic Books, Inc.

Jöreskog, Karl G. and Doug Sörbom
1978 Lisrel IV: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood. Chicago: National Educational Resources, Inc.

Kerckhoff, Alan C.
1976 "The status attainment process: socialization or allocation?" Social Forces 55:368–89.

Kerckhoff, Alan C. and Richard T. Campbell
1977a "Race and social status differences in the explanation of educational ambitions." Social Forces 55:701–14.
1977b "Black-white differences in the educational attainment process." Sociology of Education 50:15–27.

Kerlinger, Fred N. and Elazar J. Pedhazur
1973 Multiple Regression in Behavioral Research. New York: Holt, Rinehart and Winston.

Porter, James N.
1974 "Race, socialization, and mobility in educational and early occupational attainment." American Sociological Review 39:303–16.

Portes, Alejandro and Kenneth L. Wilson
1976 "Black-white differences in educational attainment." American Sociological Review 41:414–531.

Sewell, William, Archibald Haller, and Alejandro Portes
1969 "The educational and early occupational attainment process." American Sociological Review 34:82–92.

Thomas, Carl E.
1980 "Race and sex differences and similarities in the process of college entry." International Journal of Higher Education 9:179–202.

Turner, Ralph H.
1960 "Sponsored and contest mobility and the school system." American Sociological Review 25:855–67.

RACIAL TOLERANCE AS A FUNCTION OF GROUP POSITION*

A. Wade Smith
Arizona State University and *National Opinion Research Center*

White people's attitudes toward racial desegregation of the public schools are assessed by using items replicated on 22 surveys from 1954 to 1978. Myrdal's statement that racial tolerance reflects adherence to a general principle (equal access to educational opportunity) is compared to Blumer's thesis that racial tolerance varies as a function of the perceived position of groups in a racially stratified system. Depending on the racial mixture implied, whites have different levels of tolerance of school desegregation; but regardless of the racial mix, regional differences in tolerance decrease over time. Where differences in tolerance exist among other demographic subgroups of whites, those differences remain relatively constant.

Since World War II, research in race relations has tried to explore and explain individual-level racial prejudice. The basic assumptions of this work have been that individual racism forms a foundation without which structural aspects of the practice could not exist, and that at the core of the racial problems in America is the conflict between principles and/or ideals espoused by the American Creed and the norms of exclusion regarding blacks (Myrdal, 1944).

Stouffer (1955) and others have shown that white Southerners are generally less tolerant than non-Southerners and that the young and most educated tend to be the most tolerant subgroups in the population (Erskine, 1962; Schwartz, 1967; Campbell, 1971; Davis, 1975b; Williams et al., 1976). Previous research confirms similar relationships among various issues of racial tolerance and respondent's region, education, and chronological age (Hyman and Sheatsley, 1956, 1964; Sheatsley, 1965; Greeley and Sheatsley, 1971; Taylor et al., 1978; Condran, 1979).

In those studies, Southerners have shown the least tolerant attitudes on racial issues. Though Myrdal recognized that problems in race relations were not exclusive to the South, it was Greeley (1974) who demonstrated that most of the recent increase in the tolerance of whites for blacks resulted from changes in the attitudes of Southerners. There remains some disagreement as to the magnitude and long-term trends of these differences (Schwartz, 1967; Campbell, 1971; Bellingsfield, 1972; Greeley, 1974; Condran, 1979).

More recently a growing body of evidence suggests that Myrdal's predictions regarding the existence and persistence of differences in tolerance toward blacks within the white population are in error. Specifically, Myrdal erred when he foresaw a gradual erosion of age- and educational-group differences in white people's adherence to the principle of school desegregation. The predictions Myrdal made regarding the contributions of different subgroups to the general increase in tolerance over time are unsubstantiated (Smith and Moore, 1980, 1981; Smith, 1981). Finally, while Myrdal minimized the effects of different racial compositions on white attitudes toward blacks, these have been shown to influence white attitudes (Campbell, 1971; Knapp and Alston, 1972; Longshore, 1981).

Blumer's (1958, 1965) framework attempts to place these arguments in perspective—since prejudice presupposes an invidious distinction of in-group versus

* Direct all communication to A. Wade Smith, Department of Sociology, Arizona State University, Tempe, AZ 85287.

The research here was partially supported by an NSF grant (SES-8023396) to the National Opinion Research Center. The manuscript was greatly improved as a result of the helpful criticism of the ASR referees and editors. Of course, the author is solely responsible for its contents. An earlier version of this paper was presented at the meetings of the American Sociological Association, New York, 1980.

out-group on the part of the dominant members of the racial order, all feelings of prejudice are direct results of racial group membership.[1] Several factors of interracial attitudes are necessary for the development, persistence, and dispersal of prejudicial or intolerant behavior: racial identification made of one's self and others, a feeling of proprietary claim to certain areas of privilege and advantage, and a fear and/or suspicion that the subordinate race harbors designs on the prerogatives of the dominant race. These factors interact to produce a *sense of position* from which members of the dominant group view their relations with members of racial out-groups. Blumer suggests that tolerance of a racial out-group depends upon the specific mixture of in- to out-group members such tolerance produces. A racial mixture which in-group members perceive as threatening to their control over a valuable resource will extract less tolerance than a noticeably more favorable composition.

This theoretical statement is not developed precisely enough to permit construction of a quantitative model of the sequence of development of perceived intergroup competition and hostility; nor are sufficient data available to construct such a model. However, the basic insight—that structural factors such as group position and differential access to social privilege affect attitudes, rather than the opposite causal sequence of attitudes affecting structural variables such as segregation laws—is important. As noted, the causal primacy of structural factors was dismissed by Myrdal (1944) and otherwise largely overlooked by other researchers in race relations. Further, any attempt to link changes in specific components of the social structure (e.g., the desegregation of schools in specific areas) with changes in attitudes regarding participant groups will not address the initial formation and subsequent maintenance of those beliefs. The purpose here is to determine whether racial attitudes exist as a result of a generalized orientation to a set of values, as Myrdal asserted, or on the basis of a perceived position of self (interests), as Blumer suggests.

TESTING BLUMER'S AND MYRDAL'S PREDICTIONS

This paper compares the expression of beliefs in general principles of racial interaction with attitudes of intolerance, which may under specific circumstances be used to justify exclusionary practices regarding blacks. In particular, the paper examines what the difference is, if any, between tolerance for the principle of racial equality and tolerance for specific racial mixtures in a specific setting such as the schools; and what is the nature of the inconsistencies, if any, between demographic subgroups in their views of racial tolerance involving school desegregation.

Determining the Sense of Group Position

It is first necessary to identify items which reflect the sense of group position. Then a determination is possible of whether racial attitudes are influenced by the individual's perception of potential structural changes relative to his/her group's position in the social order.

Racial identification in America has not been a problem. It is generally obvious whether one is black or white. Racial identification of one's self, as well as others, is both relatively constant and something at which Americans are very adept (Quinn, 1954). It is assumed that the processes and effects of racial identification were operative throughout the span of this investigation.

The sense of proprietary claim is of crucial importance in this typology of racial tolerance. According to Blumer (1958:4) the dominant group feels that it is entitled to exclusive or prior rights in such areas as the ownership of choice real estate, access to desirable occupations, and membership in institutions such as schools.

A vital component of the group-position framework is the distribution of some scarce resource along racial lines. The controversy regarding the 1954 Brown vs.

[1] Many consider Blumer's view a derivative of Mead's discussion of the social construction of reality. At present, however, the matter is open to debate among symbolic interactionists (McPhail and Rexroat, 1979; Blumer, 1980).

Topeka Board of Education case before the U.S. Supreme Court illuminated the sense of proprietary claim to public education held by many whites. Similarly, recent decisions by the U.S. Supreme Court in the University of California–Davis Medical School vs. Bakke and Webber vs. Kaiser Aluminum (U.S. Supreme Court, 1979) cases reflect the continued emphasis on educational processes which differentially place racial group members in selected occupations and professions in American society.

Fear or apprehension that the subordinate racial group is threatening or will threaten the position of the dominant group is the final attribute necessary to the development of racial tolerance as a function of group position. This means the dominant group must perceive the subordinate group as valuing the same resource(s); the dominant group is not concerned with the subordinate group as such, but with its position relative to the subordinate group (Schermerhorn, 1956, 1965; Blumer, 1958, 1965; Glenn, 1966). Many whites feel strong competition with blacks in the desegregation and racial composition of the public schools (Campbell, 1971; Knapp and Alston, 1972; Schuman and Hatchett, 1974).

The following hypotheses form a direct approach to these issues:

Hypothesis 1: Tolerance of the principle of school desegregation and tolerance of actual school desegregation vary independently when the compositions of the racial groups in the schools vary. (As a corollary, there should be a discernible change in the level of tolerance as the proportion of blacks in the school increases.) These differences should persist over time.

Hypothesis 2: Tolerance of school desegregation differs between demographic subgroups when the racial compositions of the schools vary. These differences will persist over time.

Hypothesis 3: Differences between demographic subgroups in support of school desegregation diminish as the proportion of the racial out-group increases and the relative position of the dominant group deteriorates.

If the group-position model applies to tolerance of school desegregation, the proportion tolerant of desegregation should vary according to the perceived positions of the racial groups (i.e., the racial composition implied by a desegregated school). Blumer's group-position framework also leads to the expectation that the distribution of tolerant responses among subgroups of the population would vary according to the proposed mixture of the racial groups in the schools. If the tolerance of school desegregation varies with the relative positions of the racial groups, these differences should persist over time. Finally, subgroups of the population should manifest fewer differences in tolerance as they move from expecting to be in a majority position to expecting to occupy equal or minority positions.

This research focuses on white people's tolerance for varying degrees of school desegregation. Too few blacks are included in national surveys to permit extensive analyses of their responses. Until recently, there were no representative samples of the black population, and survey questions regarding racial tolerance were not addressed to black respondents.

DATA AND METHODS

The Data

Beginning in 1954, the American Institute of Public Opinion (AIPO) asked the survey question: "The Supreme Court has ruled that racial segregation in the public schools is illegal. This means that all children, no matter what their race, must be allowed to go to the same school. Do you approve or disapprove of this decision?" Beginning in 1964, the National Opinion Research Center (NORC) inquired of its respondents: "Do you think white students and Negro students should go to the same schools, or to separate schools?" In each case, "don't know" and "no answer" responses are excluded. This research combines the responses from replications of these questions to form a pool of surveys from 1954 through 1978.[2]

A series of survey questions assess tolerance for varying degrees of school desegregation. Since they first appeared in

[2] For a detailed empirical justification of this procedure, see Smith, 1978 and Smith, 1981.

1958, these questions have been replicated twelve times, using the same wording:[3]

1. "Would you yourself have any objection to sending your children to a school where a few of the children are (Negroes/blacks)?"
2. "Would you yourself have any objection to sending your children to a school where half of the children are (Negroes/blacks)?"
3. "Would you yourself have any objection to sending your children to a school where most of the children are (Negroes/blacks)?"

In each case the available responses were "yes," "no," "don't know," or "no answer." An affirmative response indicates intolerance, while a negative response indicates tolerance. An intolerant response to any question results in the classification of that respondent as intolerant on all subsequent questions. Thus, where it was the only item assessed, some replications of the third question in the series have been omitted. "Don't know" and "no answer" are excluded from analysis.[4]

The expression of tolerance of school desegregation, as measured by the pool combining the "court vote" and "same schools" questions, indicates adherence to the *principle* of racial equality. Desegregated situations where the school population consists of "a few blacks," "about 50% blacks," and "mostly blacks" provide data regarding tolerance, under different group positions. These varying group positions potentially place whites in positions of numerical dominance, equality, and minority in a school setting.

Independent Variables

Previous research has demonstrated that respondent's region of residence, education, and age at time of interview are significantly related to racial tolerance. For comparison, this research employs the same respondent characteristics.

Respondent's region is either Southern or non-Southern. The education of the respondent is measured by the number of years of formal schooling: less than high school (0–11 years of schooling), high-school graduate (12 years of schooling), and more than a high-school education (13 or more years of schooling including all education beyond the high-school diploma such as technical and trade schools or college). The sizes of the categories remain stable over the years covered in our survey pool.

White Americans are divided into four cohorts and assigned by year of birth. Each cohort represents a different historical context in which the respondents were raised:

1. *The Older or World War I Cohort*—born in or before 1906, these respondents were at least 48 years old in 1954 and 72 or older in 1978.
2. *The Middle or Depression Cohort*—persons age 31–47 in 1954 and 55–71 in 1978. They were born between 1907 and 1923, and reached age 21 between 1928 and 1944.
3. *The Young or World War II Cohort*—persons born between 1924 and 1937, reaching the age of majority between 1945 and 1958. These people were age 21–30 in 1954, and 39–54 in 1978.
4. *The New or Vietnam Cohort*—those born between 1938 and 1958. These persons were 18–38 in 1978, and were not included in the survey population (i.e., adults) until 1961. Since these people were too young to be included in pre-1961. Since these people were too young to be included in pre-1961 studies, calculations for this group are based on only nine surveys.

In assessing historical changes in the structure of racial attitudes, it is important to consider the "flavor" of the nation at the time a cohort is socialized. These groupings show the effects that the passage of time may have on persons born more or less in the same era (Evan, 1959; Ryder, 1965). The Appendix presents a contingency table for each dependent variable by each independent variable for the 22 surveys comprising the data pool.

---

[3] Each of the 22 surveys used in this research was drawn from modified probability samples of the noninstitutionalized population of the continental U.S. Some quotas are employed in the final stages of sampling selection.

[4] "Don't Know" and "No Answer" responses are obtained from insignificant proportions of the total sample. These proportions are relatively stable over time, with slight decreases (<1%) occurring in the later years, due to "house effects" (see Smith, 1978). The calculations here are not significantly different when "Don't Know" responses are included.

Methods and Procedures

Survey metric analysis is the procedure by which a pool of sample surveys may be examined for proportional differences (herein called "effects") between categories of background characteristics of respondents. The effects may be described using a statistic (d) stipulating either no difference, a constant difference, or a linear trend relationship between categories of background variables and respondent attitudes, over time. If respondents of different backgrounds obtain similar distributions within categories of dependent variables, this produces an effect of no difference. If categories of variables maintain differences in proportionate replies, these differences are said to be constant or consistent. Finally, should significant differences exist between categories, but these differences either increase or decrease over time, the effects are presented as linear or quasi-linear models (Davis, 1974, 1975a, 1976).

At each stage of the analysis a goodness-of-fit test, using a chi square adjusted for multistage samples, is employed.[5] This adjusted chi square indicates whether a hypothesized model accurately represents the data. If a hypothesis is rejected, the next step is to explore a more advanced model. Although the no-difference model may appear to adequately represent the data, it is sometimes advantageous to test the accuracy of the description of a more advanced model (e.g., of a constant difference) so that otherwise nonsalient relationships can be explored (Taylor, 1980).

The model under investigation can be compared to prior models to determine if any significant improvement in the description of the data has been obtained. For example, by subtracting the chi square (and degrees of freedom) used in the test of the constant difference hypothesis from the chi square (and corresponding degrees of freedom) obtained in the test of no difference, the result is a one degree of freedom test of the improvement of the latter model over the former. The accuracy of the description of a linear model may be similarly compared to that offered by the constant difference model. If the resulting goodness-of-fit is not significant, then the extent to which a more advanced model improves the description of the data is borderline.[6]

FINDINGS

If racial tolerance reflects an adherence to a generalized principle of equal access to resources, then tolerance for school desegregation in principle should equal, or at least approximate, tolerance for all of the varying degrees of actual school desegregation. Further, any initial differences between tolerance in principle and tolerance for specific racial mixtures is expected to disappear over time, as desegregation becomes the norm rather than the exception. This is not the case (Figure 1). Between 1954 and 1978, there remain substantial differences between tolerance of school desegregation in principle and white people's tolerance for every potential mixture in a desegregated school.

Majority white schools—those with only a "few" blacks—are the least objectionable to whites. Since 1958, whites have been more tolerant of majority white schools than any other type. The principle of school desegregation obtains less acceptance than majority white schools. Schools whose populations are half white and half black consistently receive less acceptance than either majority white schools or desegregation in principle. Majority black schools receive the least tolerance from whites.

[5] Since all of the samples used in this analysis were constructed in two stages (the selection of households and the quotas on respondent characteristics), it is necessary to adjust the chi square, multiplying it by 1.414 (Davis, 1976). The degrees of freedom are determined by the number of surveys in the data pool. Each succeeding model consumes one degree of freedom in its test (e.g., the no difference model tests with 15 d.f., the constant difference model tests with 14 d.f., etc.).

[6] While other methodologies may appear to be more applicable, survey metric procedures yield superior results given the research questions of interest here, while minimizing the potential for the types of errors associated with several parametric techniques increasingly used in contingency table analyses. Davis (1976) compares survey metric procedures with other techniques; and Swafford (1980) notes the vulnerabilities of other methodologies.

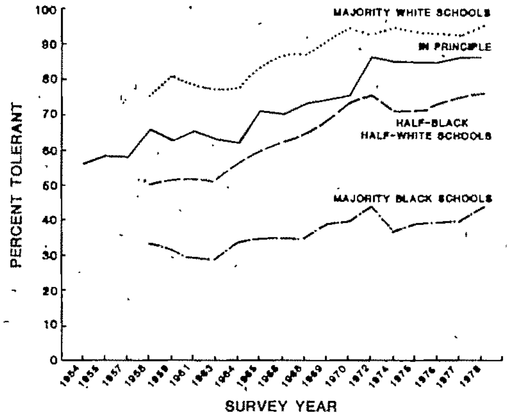


Figure 1. Tolerance of Varying Degrees of School Desegregation 1954–1978.

While the general ordering of the level of tolerance for the school desegregation items is static over time, there are some changes in the level of tolerance for each item (summarized in Table 1). While tolerance of desegregation in principle becomes more like tolerance of majority white schools (by 10.4%), tolerance for half-black and majority-black schools becomes less similar to tolerance for the principle of desegregated schools (by 4.0% and 19.9%, respectively). However, neither the total percent change in tolerance nor the percent change per year occurs in the relationship expected if the group-position framework were to apply.

The existing differences in the level of tolerance for school desegregation vary according to the extent of desegregation implied, and these differences persist and even increase over time. Contrary to Myrdal's inference, acceptance of each of the implied racial mixtures in a desegregated school is unlike the acceptance of desegregation in principle. Yet, if racial tolerance is a function of group position as Blumer asserts, there should be differences within the white population in their tolerance for varying degrees of school desegregation. Where they exist, these differences should persist over time, and there should be fewer differences (i.e., greater agreement) among whites as their potential group position within a school moves from numerical dominance to subordination.

Tolerance of the Principle of School Desegregation

Previous research has demonstrated that subgroup differences exist regarding white people's tolerance for school desegregation in principle. Regional differences were found to decrease at an average rate of 1.47% per year. This rate of convergence is not strictly linear, but does contain a substantial linear component. On the other hand, cohort and educational subgroup differences were consistent over time. The youngest (Vietnam) cohort favored school desegregation 7.0% more than the Depression cohort (base group), while the World War II cohort was similarly 2.5% more tolerant throughout the span of the survey pool. The World War I cohort was significantly less tolerant (8.2%) than the base group within this same time frame. Respondents with less than a high-school education were 16.4% less tolerant than high-school graduates, while those at least with some college education demonstrated 6.9% more tolerance on this issue over time. (See Smith (1981) for a fuller discussion of the procedures involved in obtaining particular effects and of the contributions of subgroup differences toward the overall rate of change in tolerance of the principle of school desegregation.)

Table 1. Tolerance of Varying Degrees of School Desegregation, 1954–1978, Total White Respondents

Type of desegregation	Initial tolerance level		Final tolerance level		Percent change in tolerance	Percent change per year
In principle	55.8%	(1954)[a]	86.2%	(1977)	+30.4	1.32
Majority white schools	75.2	(1958)	95.2	(1978)	+20.0	1.00
Half-white, half-black schools	49.8	(1958)	76.2	(1978)	+26.4	1.32
Majority black schools	33.1	(1958)	43.6	(1978)	+10.5	.53

[a] There were two surveys in 1954 containing this question. This percentage is their mean observed level of tolerance. Survey year appears in parentheses.

Tolerance of Majority White Schools

Some differences persist among demographic subgroups of whites in their tolerance of majority white schools (see Table 2). Regional differences in tolerance do exist, yet they tend to diminish at a linear or quasi-linear rate. While the linear model does offer a significant improvement over previous hypotheses, caution would dictate that the 2.33% dissipation in regional differences be considered a weighted average of the change over time. Although not strictly linear, this rate of change is best described using a linear equation. At the least measurement, there remained nonzero differences between regions in their tolerance of majority white but desegregated schools.

Except for the Vietnam cohort, there are no significant age-group differences regarding white people's tolerance for majority white schools (Table 2). This newest cohort averages 4.2% more tolerance than the Depression cohort in their tolerance of sending their children to

Table 2. Tolerance of Majority White Schools, 1958–1978

	Hypothesis	Adjusted Chi Square	Degrees of Freedom	Probability	Decision
Summary Model of Regional Differences (Base = Northern respondents)					
South	No difference	500.7	12	<.001	Reject
	Constant difference	292.5	11	<.001	Reject
	Linear trend	89.7	10	<.001	Reject
	Reduction from the linear parameter	202.8	1	<.001	Significant
	Final Model[a] d = −.564 + 0.0233 (Year − 1958) (Plus interaction terms in later years)				
Summary Model of Cohort Differences (Base = Middle cohort)					
New cohort	No difference	21.1	10	.021	Reject
	Constant difference	3.8	9	.927	Accept
	Reduction from the constant parameter	17.3	1	<.001	Significant
Young cohort	No difference	17.4	12	.134	Accept
Old cohort	No difference	15.8	12	.201	Accept
	Final Model[b]				
	New cohort	d = .042	σ = .010		
	Young cohort	d = 0			
	Old cohort	d = 0			
Summary Model of Educational differences (Base = High school graduates)					
Less than high-school graduate	No difference	120.0	12	<.001	Reject
	Constant difference	18.9	11	.062	Accept
	Reduction from constant parameter	101.1	1	<.001	Significant
Some college	No difference	9.6	12	.654	Accept
	Constant difference	5.7	11	.894	Accept
	Reduction from the constant parameter	3.9	1	.05	Significant
	Final Model[c]				
	Less than high-school graduate		d = −.092	σ = .009	
	Some college		d = .014	σ = .007	

NOTE: σ denotes standard error of the estimate.

[a] In its tolerance of majority white schools, the South becomes more like the non-South at a rate of about 2.33% per year.

[b] In their tolerance of a few blacks in school with their children, the new cohort is 4.2% more tolerant of majority white schools than the rest of the population.

[c] Respondents with less than a high-school education are 9.2% less tolerant of majority white schools than high-school graduates. Respondents with some college education are 1.4% more tolerant than those with only a high-school education.

schools which have "a few" blacks. Since no other cohorts differ from the base group, the Vietnam cohort is effectively 4.2% more tolerant than all others on this issue.

Consistent differences exist between whites of different educational backgrounds in their tolerance for desegregated, majority white schools (Table 2). The least educated whites average 9.2% less tolerance than high-school graduates. The most educated segment of the white population consistently averages 1.4% more tolerance regarding sending their children to schools having a few blacks. In short, the most educated whites are 10.6% more tolerant of majority white schools than the least educated.

Tolerance of Half-Black Schools

Demographic differences in white people's tolerance of half-black schools is similar to their tolerance of minimally integrated schools (see Table 3). Regional differences in white people's objections to

Table 3. Tolerance of Half-Black Half-White Schools, 1958–1978

	Hypothesis	Adjusted Chi Square	Degrees of Freedom	Probability	Decision
	Summary Model of Regional Differences (Base = Northern Respondents)				
South	No difference	427.6	12	<.001	Reject
	Constant difference	147.5	11	<.001	Reject
	Linear trend	41.8	10	<.001	Reject
	Reduction from the linear parameter	105.7	1	<.001	Significant
	Final Model[a]				
	South $d = -.411 + 0.0202$ (Year − 1958)				
	(Plus interaction terms in middle years)				
	Summary Model of Cohort Differences (Base = Middle cohort)				
New cohort	No difference	39.3	10	<.001	Reject
	Constant difference	7.5	9	.584	Accept
	Reduction from the constant parameter	31.8	1	<.001	Significant
Young cohort	No difference	10.5	12	.570	Accept
Old cohort	No difference	12.8	12	.387	Accept
	Final Model[b]				
	New cohort	$d = 0.106$	$\sigma = .019$		
	Young cohort	$d = 0$			
	Old cohort	$d = 0$			
	Summary Model of Educational Effects (Base = High-school graduates)				
Less than high-school graduate	No difference	53.3	12	<.001	Reject
	Constant difference	11.4	11	.412	Accept
	Reduction from the constant parameter	41.9	1	<.001	Significant
Some college	No difference	11.7	12	.529	Reject
	Constant difference	6.6	10	.833	Accept
	Reduction from the constant parameter	5.1	1	<.05	Significant
	Final Model[c]				
	Less than high-school graduate	$d = -.091$	$\sigma = .014$		
	Some college	$d = .033$	$\sigma = .015$		

NOTE: σ denotes standard error of the estimate.

[a] Differences in levels of tolerance between the South and North are disappearing at the rate of 2.02% per year on this issue.

[b] The new cohort is a consistent 10.6% more tolerant than other cohorts on this issue.

[c] Respondents with less than a high-school diploma are constantly 9.1% less tolerant of equally desegregated schools than high-school graduates. Respondents having some college education are consistently 3.3% more tolerant than high-school graduates on this issue.

sending their children to a school whose composition is 50% black, diminish at about 2.02% per year. Though the non-South is no longer 41.1% more tolerant than the South on this issue, a small difference remains between regions in their tolerance of equally integrated schools. Again, cohort differences obtain for the newest age group. They appear 10.6% more tolerant on this issue than the Depression cohort.

With respect to education, whites having less than a high-school education appear 9.1% less tolerant than the base group—high-school graduates (Table 3). Those educated beyond high school are 3.3% more tolerant of half-black schools than the base group. Arithmetically, the most educated whites are 12.4% more tolerant of half-black, half-white schools than the least educated.

Tolerance of Majority Black Schools

Regional differences continue to exist in white people's tolerance of majority black schools (see Table 4). While it is just an average, these differences diminish at a rate of approximately 1.44% per year. A small difference remains between Southern and non-Southern whites in their tolerance of majority black schools.

On this issue, the Vietnam cohort is again the most tolerant age group. Although it might at first appear that they are no different from others, the significant reduction from the constant term warrants consideration of their consistent 5.5%

Table 4. Tolerance of Majority Black Schools 1958–1978

	Hypothesis	Adjusted Chi Square	Degrees of Freedom	Probability	Decision
	Summary Model of Regional Differences (Base = Northern Respondents)				
South	No difference	249.5	12	<.001	Reject
	Constant difference	63.5	11	<.001	Reject
	Linear trend	27.5	10	<.001	Reject
	Reduction from the linear parameter	36.0	1	<.001	Significant
	Final Model[a] South d = −.285 + 0.0124 (Year − 1958) (Plus interactions in later years)				
	Summary Model of Cohort Differences (Base = Middle cohort)				
New cohort	No difference	12.9	10	.228	Accept
	Constant difference	6.4	9	.696	Accept
	Reduction from the constant parameter	6.5	1	<.05	Significant
Young cohort	No difference	15.8	12	.201	Accept
Old cohort	No difference	6.3	12	.903	Accept
	Final Model[b]				
	New	d = .055	σ = .021		
	Young	d = 0			
	Old	d = 0			
	Summary Model of Educational Differences (Base = High-school graduates)				
Less than high-school graduate	No difference	9.1	12	.700	Accept
Some college	No difference	8.5	12	.747	Accept
	Final Model[c]				
	Less than high-school graduate		d = 0		
	Some college		d = 0		

NOTE: σ denotes standard error of the estimate.

[a] The difference in tolerance of majority black schools beween the South and the North is disappearing at the rate of 1.44% per year.

[b] Regarding white people's tolerance of majority black schools, the new cohort is consistently 5.5% more tolerant than others.

[c] Regarding their tolerance of predominantly black schools, white people have no difference by educational attainment.

differences in tolerance. The young and old cohorts fail to demonstrate differences between themselves and the middle cohort. There are no educational-group differences in the tolerance of majority black schools.

Cohort, Education, and Racial Tolerance

The results presented thus far reflect only the bivariate relationships (over time) between region, cohort, education, and racial tolerance. Yet there are known interrelationships among the independent variables. Cohort membership and educational attainment are highly correlated, and an interaction exists between region, education, and time.[7] The question arises as to whether the bivariate relationships explored to this point persist when other independent variables are incorporated into the model.

Briefly, the consideration of multivariate models for tolerance of the principle of desegregation, majority white, and majority black desegregated schools is omitted. Desegregation in principle has proven to be an inexact measure, since greater tolerance is obtained for majority white schools, and the extent of changes in tolerance among demographic groups is unlike any other measure. Of the varying degrees of school desegregation, tolerance of majority white schools changes the most over time, and attitudes toward majority black schools change the least; so a test involving the median issue is offered as a compromise.

The tests of the various multivariate models are presented in Table 5. As in the previous tables, the relationships between region, cohort, education, and tolerance over time are presented; however, in this case the remaining independent variables are held constant. The creation of additional cells in the table increases the degree of freedom, making it more difficult to obtain effects on the basis of consistency over time and giving greater weight to the magnitude of any categorical differences.

There is little change in the type or magnitude of effects obtained in earlier analyses. When cohort and education are controlled, the South converges on the non-South in tolerance of half-black schools. The rate of convergence is somewhat slower than before, but the difference of .61% per year (from Table 3) is trivial. Over time, the Vietnam cohort remains the most tolerant on this issue, even when region and education are controlled. However, here the Vietnam cohort is only 8.8% more tolerant than all others, as opposed to the 10.6% difference obtained earlier. Educational-group differences over time persist on this issue when region and cohort membership are controlled. The least educated are 5.4% less tolerant than the base group (down from the 9.1% difference obtained earlier). However, the college educated fail to obtain any difference from high-school graduates. The small (3.3%) difference manifested earlier by the consistency of their responses on this issue does not exist when region and cohort are controlled. Multivariate models of racial tolerance over time offer very small differences when compared to bivariate models. This suggests that the variables of region, cohort membership, and educational attainment operate relatively independently in their effects on changes in the level of tolerance of school desegregation.

Overall, the evidence can best be evaluated by addressing the specific hypotheses as they were formulated earlier.

Hypothesis 1. Confirmed: Tolerance of the principle of school desegregation is not obtained in the same proportions as tolerance for any of the specific mixtures of desegregated schools. Further, as the respondent's racial group moved from numerical dominance, through equality, to numerical minority, decreasing levels of racial tolerance were recorded. Although all measures of tolerance of school desegregation obtained some increases over time, there was no change in the relative order of approval.

Hypothesis 2. Confirmed: Demographic subgroups of whites display, over time, differences in tolerance on each of the measures of varying degrees of school desegregation. Regional differences exist

---

[7] Southern high-school graduates increasingly attend college in greater proportions than their non-Southern counterparts (Davis, 1975a).

Table 5. Tolerance of Half-Black Half-White Schools, 1958–1978

	Hypothesis	Adjusted Chi Square	Degrees of Freedom	Probability	Decision
	Summary Model of Regional Differences Controlling for Cohort and Education (Base = Northern Respondents)				
South	No difference	557.0	144	<.001	Reject
	Constant difference	265.5	143	<.001	Reject
	Linear trend	79.8	142	<.001	Reject
	Reduction from the linear parameter	185.7	1	<.001	Significant
	Final Model[a] South d = −.417 + 0.0141 (Year − 1958) (Plus interaction terms in middle years)				
	Summary Model of Cohort Differences Controlling for Region and Education (Base = Middle cohort)				
New cohort	No difference	53.6	60	<.721	Reject
	Constant difference	29.5	59	.955	Accept
	Reduction from the constant parameter	24.1	1	<.001	Significant
Young cohort	No difference	36.7	72	1.000	Accept
Old cohort	No difference	41.8	72	.387	Accept
	Final Model[b]				
	New cohort	d = 0.088	σ = .018		
	Young cohort	d = 0			
	Old cohort	d = 0			
	Summary Model of Educational Effects Controlling for Region and Cohort (Base = High-school graduates)				
Less than high-school graduate	No difference	82.0	96	.844	Accept
	Constant difference	67.9	95	1.000	Accept
	Reduction from constant parameter	11.1	1	<.001	Significant
Some college	No difference	47.8	96	1.000	Accept
	Final Model[c]				
	Less than high-school graduate	d = −.054	σ = .014		
	Some college	d = 0			

NOTE: σ denotes standard error of the estimate.

[a] Differences in levels of tolerance between the South and North are disappearing at the rate of 1.41% per year on this issue.

[b] The new cohort is a consistent 8.8% more tolerant than other cohorts on this issue.

[c] Respondents with less than a high-school diploma are constantly 5.4% less tolerant of equally desegregated schools than high-school graduates. Respondents having some college education are no more tolerant than high-school graduates on this issue.

and dissipate at linear or quasi-linear rates over time on all issues. Consistent cohort differences in tolerance are obtained over time regardless of the racial mixture implied by school desegregation. Only majority black schools fail to differentiate educational subgroups over time.

Hypothesis 3. Partially confirmed: While attitudes toward a majority black school do elicit greater agreement among white subgroups than a majority white school, there is not always an increase in agreement as their group position diminishes. As the issue shifts from majority white schools, to equality of numbers, to majority black schools, the annual rate at which regional differences converge decreases (2.33%, 2.02%, 1.24%), while cohort differences increase and then decrease (4.2%, 10.6%, 5.5%). The difference in tolerance between the most and least educated whites is 10.6% for majority white desegregated schools, increases to 12.4% for half-black, half-white schools, and disappears when the issue is tolerance of mostly black schools. Although the disappearance of educational-group differences as their group position diminishes is not consistent, whites manifest the greatest agreement in their toler-

ance when desegregation places them in the minority.

CONCLUSIONS AND IMPLICATIONS

In the final analysis, these are the results expected if the group-position framework applies to white people's tolerance of school desegregation. The level of tolerance manifested by whites does diminish as the potential proportion of blacks in a desegregated school increases. Not only does the level of tolerance vary according to the extent of the desegregation involved, but where they exist, differences in tolerance among demographic subgroups persist over time. There is some decline in regional differences, but to this point nonzero discrepancies remain. However, as their position in the racial order moves from a numerical majority to a numerical minority, there is a greater agreement among whites in their tolerance for school desegregation. Further, regardless of the racial mixture implied by desegregation, regional, cohort, and educational subgroups of whites display patterns of tolerance markedly dissimilar to their tolerance of school desegregation in principle. The evidence of response patterns by the total white population, as well as selected demographic subgroups, supports the thesis that racial tolerance exists as a function of the sense of group position.

The sense of proprietary claim—the objections of whites when many blacks come into "their" schools—identified in Blumer's statement is apparent here in that whites of all regional, cohort, and educational attainment groups share a common self-interest in their unwillingness to accept minority dominance. The regional trends converge more slowly for tolerance of majority black schools than for tolerance of desegregation in principle, majority white, or half-black, half-white schools. The cohort differences are smaller regarding tolerance of majority black schools than for desegregation in principle or half-black schools. Finally, educational-group differences are smaller in tolerance of majority black schools than on all other issues. Eight of the nine comparisons work in the direction expected by Blumer's framework. While whites of non-Southern origin, or born in the youngest cohort, or among the most educated may be more tolerant than the Southern, oldest, or least educated, all are wary of inundation by blacks.

Since tolerance of school desegregation in principle is not the same as tolerance of any specific measure of school desegregation, there is a danger of attributing too much importance to (and confusing interpretations of) the exact magnitude of the differences between, say, tolerance in principle and tolerance of mostly-white integrated schools. The findings here may imply that, over time, a small proportion of the white population rejects school desegregation in principle and yet has no objection to sending their children to schools enrolling a few blacks (see Table 1 and Figure 1). An equally appealing interpretation is that some whites reject school desegregation in principle because the relative proportion of blacks in a potentially desegregated school is unknown, but favor desegregation when the group position is specified in their favor. This would explain why, of all the items, tolerance of the principle of desegregation experiences the most increase over time. Many whites came to realize that in a desegregated school, they would continue to be the majority group.

These findings suggest a source and solution to cognitive inconsistencies first observed by Myrdal. In the past, individuals who held unegalitarian values regarding racial others were viewed as "prejudiced." In contrast to the stated egalitarian nature of a society of potentially unlimited mobility, the result seemed like an internal dilemma. The data here indicate that while individuals may now hold general egalitarian values (i.e., they are free of "old fashioned" bigotry), they continue to manifest attitudes about specific areas of life that are incongruous with these ideals. The result is conflict between societal goals and the conduct of individuals when they have the opportunity to operationalize these ideals. In fact, there is no internal dilemma, because individuals predetermine a set of conditions under which they will adhere to their principles. If these conditions do not exist, the ideals become inoperative.

APPENDIX Percent Tolerant of Varying Degrees of School Desegregation, 1954–1978

Tolerance Measure of:		Survey Year																					
		1954	1954[a]	1955	1955[a]	1957	1958	1959	1961	1963	1964	1965	1965[a]	1966	1968	1969	1970	1972	1974	1975	1976	1977	1978
Desegregation in principle[b]																							
Region	NORTH	69.0	68.3	75.8	72.0	74.1	76.5	75.4	78.1	*	74.2	81.1	*	78.2	84.0	*	86.8	92.1	*	*	91.7	91.9	*
	SOUTH	23.5	19.5	16.0	16.9	16.9	32.9	24.3	27.7		28.3	43.1		49.0	45.0		47.2	68.3			68.9	72.8	
Cohort	NEW	—[d]	—	—	—	—	—	—	72.5	*	64.6	72.3	*	69.9	78.5	*	85.5	94.1	*	*	90.3	90.9	*
	YOUNG	63.2	58.1	58.7	61.2	63.6	70.6	69.1	67.0		65.3	76.8		74.7	75.0		77.9	85.7			83.4	88.1	
	MIDDLE	56.9	56.5	61.4	61.2	55.6	65.3	65.1	68.1		65.0	71.4		69.8	72.6		74.5	84.1			83.8	82.3	
	OLD	50.2	50.8	57.0	51.3	55.3	61.9	53.9	59.0		51.0	60.8		62.8	67.6		65.3	74.4			70.0	69.0	
Education	LHSG	45.9	45.6	51.4	48.1	48.9	55.2	51.0	56.3	*	48.6	57.5	*	61.7	62.8	*	61.3	73.5	*	*	73.9	74.5	*
	HSG	66.4	59.6	66.4	64.0	65.3	70.0	69.5	72.2		66.6	74.8		74.4	76.3		82.1	91.1			87.7	90.1	
	HSG+	74.9	74.2	71.0	69.8	68.7	82.6	75.5	77.6		79.4	84.9		78.4	88.6		90.2	96.1			92.9	95.5	
Total		56.6	54.9	58.8	57.5	57.7	65.7	62.3	65.0	*	61.7	71.0	*	70.0	73.3	*	75.5	86.1	*	*	84.7	86.2	*
Majority white schools																							
Region	NORTH	*	*	*	*	*	88.6	93.5	*	89.5	*	92.4	92.9	*	*	90.1	94.6	96.0	97.4	94.1	*	95.2	96.8
	SOUTH						32.4	32.4		45.3		68.8	56.5			93.0	94.2	83.3	89.5	90.8		87.0	91.6
Cohort	NEW	*	*	*	*	*	—	—	*	63.6	*	96.7	85.7	*	*	90.9	98.3	96.6	96.5	96.6	*	94.1	96.8
	YOUNG						74.9	85.2		74.7		83.7	78.6			87.7	92.7	94.3	95.6	94.3		92.8	94.3
	MIDDLE						75.7	80.1		79.4		89.9	85.2			90.6	92.9	90.2	93.1	91.0		92.2	93.8
	OLD						73.8	71.1		70.4		73.5	97.1			80.0	80.0	85.4	92.8	82.6		87.4	92.7
Education	LHSG	*	*	*	*	*	66.8	70.2	*	66.0	*	78.9	74.9	*	*	84.8	98.0	87.3	90.0	89.2	*	86.5	91.9
	HSG						82.0	87.4		84.0		90.3	84.2			89.1	95.0	95.1	96.6	95.8		95.4	96.5
	HSG+						85.7	91.3		89.2		93.5	89.2			92.4	95.1	97.1	97.9	94.0		96.8	97.1
Total		*	*	*	*	*	75.2	81.1	*	76.7	*	86.2	81.2	*	*	91.2	94.4	92.8	94.9	93.1	*	92.7	95.2
Half-black half-white schools																							
Region	NORTH	*	*	*	*	*	60.1	65.1	*	63.2	*	70.1	70.8	*	*	66.0	72.6	81.3	74.3	73.2	*	78.9	78.4
	SOUTH						19.0	16.1		23.6		36.7	32.4			70.7	75.2	59.2	63.8	66.7		66.2	71.3
Cohort	NEW	*	*	*	*	*	—	—	*	63.6	*	88.9	66.7	*	*	73.0	74.5	85.2	77.1	79.4	*	79.5	77.4
	YOUNG						51.7	53.4		54.1		63.1	62.3			63.2	74.2	71.9	65.9	64.9		74.9	78.6
	MIDDLE						49.7	58.0		48.9		59.6	55.2			61.0	71.0	77.2	68.5	69.9		70.2	73.6
	OLD						43.1	61.0		44.4		45.2	81.2			20.0	60.0	61.3	68.7	63.6		68.4	73.5
Education	LHSG	*	*	*	*	*	45.4	47.8	*	41.9	*	51.8	53.7	*	*	56.5	63.8	67.4	63.5	67.7	*	67.1	76.1
	HSG						55.1	56.2		56.2		66.7	61.3			68.0	77.4	78.3	73.1	71.1		78.3	75.3
	HSG+						53.1	67.2		66.4		67.4	61.0			67.8	72.2	84.4	76.0	76.2		80.4	77.4
Total		*	*	*	*	*	50.2	51.5	*	51.4	*	61.2	58.2	*	*	67.8	73.5	75.6	71.0	71.1	*	75.1	76.2

APPENDIX Continued

Tolerance Measure of:		1954	1954[a]	1955	1955[a]	1957	1958	1959	1961	1963	1964	1965	1965[a]	1966	1968	1969	1970	1972	1974	1975	1976	1977	1978
Majority black schools																							
Region	NORTH	*	*	*	*	*	40.1	37.7	*	37.1	*	40.7	42.4	*	*	38.1	36.4	47.7	40.3	41.0	*	42.0	46.7
	SOUTH						11.6	8.7		9.0		23.8	15.7			39.8	45.9	32.3	29.5	34.8		34.6	36.5
Cohort	NEW	*	*	*	*	*	—	—	*	45.5	*	44.4	33.3	*	*	45.0	38.8	50.0	39.9	41.2	*	44.9	44.2
	YOUNG						35.3	32.7		28.2		35.0	28.2			33.1	34.4	42.1	34.4	38.0		39.7	41.9
	MIDDLE						32.8	31.3		27.9		37.3	41.3			41.7	45.1	41.7	34.0	38.6		31.8	45.7
	OLD						29.8	32.4		41.9		32.3	45.2			20.0	40.0	39.3	39.5	37.5		38.9	43.2
Education	LHSG	*	*	*	*	*	31.0	26.4	*	22.5	*	34.7	34.4	*	*	34.2	36.6	42.1	36.4	39.2	*	38.1	43.8
	HSG						36.8	30.9		32.9		36.5	35.1			37.6	38.8	42.3	34.9	36.7		37.3	43.8
	HSG+						30.0	42.9		36.3		35.2	31.9			40.9	35.5	48.2	39.4	42.4		40.7	43.2
Total		*	*	*	*	*	33.1	31.5	*	28.7	*	36.1	33.4	*	*	38.8	39.8	43.7	36.9	39.0	*	39.8	43.6
Sample Size[e]	N =	1242	1257	1236	1232	1300	1316	1326	3083	1392	1669	1243	1122	1239	1204	547	1206	1320	1281	1300	1323	1309	1337

NOTE: For those who wish to reanalyze these data by other methods, for each category of an independent variable the level of intolerance may be obtained by subtracting the data given here from 100%. Sample sizes appear below.

[a] There were two surveys in each year of 1954, 1955, and 1965 in this data pool.

[b] Derived from Smith, 1981.

[c] This issue not surveyed during this year.

[d] The new cohort does not exist until 1961.

[e] Excludes nonwhites, do not know, and no answer responses.

REFERENCES

Bellingsfield, Gwen
1972–73 "Current research: white attitudes toward racial integration and the urban riots of the 1960s." Public Opinion Quarterly 36:579–84.

Blumer, H.
1958 "Race prejudice as a sense of group position." Pacific Sociological Review I: 3–7.
1965 "Industrialization and race relations." Pp. 220–53 in Guy Hunter (ed.), Industrialization and Race Relations. London: Oxford University Press.
1980 "Mead and Blumer: the convergent methodological prespectives of social behaviorism and symbolic interactionism." American Sociological Review 45:409–19.

Campbell, Angus
1971 White Attitudes Toward Black People. Ann Arbor, Michigan: Institute for Social Research, University of Michigan.

Condran, John G.
1979 "Changes in white attitudes towards blacks, 1963–1977." Public Opinion Quarterly 43:463–76.

Davis, James A.
1974 "Concepts and procedures for constructing odds ratio models for causally ordered, multivariate contingency tables." Pp. 1–34 in Studies in Social Change Since 1948: Volume I. Chicago: National Opinion Research Center Report 127A.
1975a "Background characteristics of the U.S. adult population, 1952–1972, a survey metric model." Pp. 1–36 in Studies in Social Change Since 1948: Volume II. Chicago: National Opinion Research Center Report 127B.
1975b "Communism, conformity, cohorts, and categories: American tolerance in 1954 and 1972–73." American Journal of Sociology 81:491–513.
1976 "D systems: analyzing contingency tables with linear flow graphs." Pp. 111–45 in David R. Heise (ed.), Sociological Methodology, San Francisco: Jossey-Bass.

Erskine, Hazel
1962 "The polls: race relations." Public Opinion Quarterly 26:178–88.

Evan, W. M.
1959 "Cohort analysis of survey data." Public Opinion Quarterly 38:63–72.

Glenn, N. D.
1966 "White gains from negro subordination." Social Problems 14:159–78.

Greeley, A. M.
1974 "Attitudes towards racial integration: the South 'catches up'." Pp. 241–50 in Lee Rainwater (ed.), Social Problems and Public Policy: Inequality and Justice. Chicago: Aldine.

Greeley, A. M. and Paul B. Sheatsley
1971 "Attitudes toward desegregation." Scientific American (December):13–19.

Hyman, H. H. and Paul B. Sheatsley
1956 "White attitudes toward desegregation." Scientific American 195 (December):35–9.
1964 "Attitudes toward desegregation." Scientific American 211 (July):16–23.

Knapp, Melvin J. and John P. Alston
1972–73 "White parental acceptance of varying degrees of school desegregation, 1965 and 1970." Public Opinion Quarterly 36:558–91.

Longshore, Douglas
1981 "The control threat in desegregated schools: exploring the relationship between school composition and intergroup hostility." Paper read at the 1981 meeting of the American Sociological Association.

McPhail, C. and C. Rexroat
1979 "Mead vs. Blumer: the divergent methodological perspectives of social behavioralism and symbolic interactionism." American Sociological Review 44:449–67.

Myrdal, Gunnar
1944 An American Dilemma. New York: Harper and Row.

Quinn, Olive Westbrook
1954 "The transmission of racial attitudes among white Southerners." Social Forces 33:41–7.

Ryder, N. B.
1965 "The cohort as a concept in the study of social change." American Sociological Review 30:843–61.

Schermerhorn, R. A.
1956 "Power as a primary concept in the study of minorities." Social Forces 35:53–6.
1965 "Toward a general theory of minority groups." Phylon 25:238–46.

Schuman, Howard and Shirley Hatchett
1974 Black Racial Attitudes: Trends and Complexities. Ann Arbor, Michigan: Institute for Social Research, University of Michigan.

Schwartz, Mildred A.
1967 Trends in White Attitudes Towards Negroes. Chicago: National Opinion Research Center, University of Chicago.

Sheatsley, Paul B.
1965 "White attitudes toward the negro." Daedalus (Winter).

Smith, A. Wade
1981 "Tolerance of school desegregation, 1954–1977," Social Forces 59:1256–74.

Smith, A. Wade and Elsie G. J. Moore
1980 "White attitudes toward school desegregation." Paper read at the 1980 meeting of the American Sociological Association.
1981 "Continuing trends in white attitudes toward school desegregation, 1954–1980." Paper read at the 1981 meeting of the American Sociological Association.

Smith, Thomas E.
1978 "In search of house effects: a comparison of responses to various questions by different survey organizations." Public Opinion Quarterly 42 (winter):443–63.

Stouffer, Samuel A.
1955 Communism, Conformity and Civil Liberties. New York: Wiley.

Sudman, Seymour and Norman Bradburn
1974 Response Effects in Surveys. Chicago: Aldine.
Swafford, Michael
1980 "Three parametric techniques for contingency table analysis: a nontechnical commentary." American Sociological Review 45:664–90.
Taylor, Garth
1980 "Procedures for evaluating trends in public opinion." Public Opinion Quarterly:86–100.
Taylor, Garth, Paul B. Sheatsley, and Andrew M. Greeley
1978 "Attitudes toward desegregation." Scientific American 238:42–49.
United States Supreme Court
1979 "1979 Abstract Sheets." U.S. Supreme Court Reports. Washington, DC: U.S. Government Printing Office.
Williams, J. A., C. Z. Nunn, and L. St. Peter
1976 "Origins of tolerance: findings from a replication of Stouffer's 'Communism, conformity, and civil liberties'." Social Forces 55:394–408.

SOCIAL MOBILITY AND FERTILITY: TWO EFFECTS IN ONE*

GILLIAN STEVENS
University of Wisconsin-Madison

None of the numerous variants of the social mobility and fertility hypothesis, each of which presumes only one effect of mobility on fertility, has received empirical support. Partitioning mobility into relative (circulation) mobility and mean (structural) mobility reconciles some of the contrary predictions. Analysis of the Occupational Changes in a Generation Study data confirms that the relationship between total mobility and fertility is weak and not always significant. However, when total mobility is partitioned into relative mobility and mean mobility, the results suggest that fertility behavior is consistently negatively related to mean mobility and positively related to relative mobility.

A relationship between social mobility and fertility was first advanced around the turn of the century when Dumont drew an analogy between the social movement of couples and the rise in a column of liquid: "just as a column of liquid has to be thin in order to rise under the force of capillarity, so a family must be small in order to rise in the social scale" (cited in Westoff, 1953). Since then, the social mobility and fertility hypothesis has undergone numerous transformations. It has been worded to refer to intergenerational mobility (Blau and Duncan, 1967), intragenerational mobility (Berent, 1952), and occupational role performance (Hargens et al., 1978). At various times it has predicted that the effects of mobility in one generation are evident in the fertility of the succeeding generation (the Galton-Fisher version), or that the effects are confined to the generation experiencing the mobility (Featherman, 1973). Even the direction of the predicted outcomes has changed. Fertility has been posited to vary inversely with mobility (Blau and Duncan, 1967), directly with mobility (Easterlin, 1973), and inversely with upward mobility and directly with downward mobility (Boyd, 1973).

In spite of the range of variation in the direction and timing of the hypothesized effects of social mobility on fertility, none of these variations has been corroborated. Each of these variations, however, presumes only one effect of mobility on fertility. Mobility experiences, however, are attributable to both structural factors (labor-market conditions, for example)

* Direct all correspondence to Gillian Stevens, Department of Sociology, University of Wisconsin, Madison, WI 53706.

A version of this paper was read at the Population Association of America meetings in Washington, D.C., April 1981. Computing services were provided by the Center for Demography and Ecology, University of Wisconsin, Madison, funded by the National Institute of Child Health and Human Development (Grant #HD-5876). I thank Larry Bumpass, Ronald Kessler, Robert Mare, Jim McGonigle, and Judith Seltzer for helpful comments on earlier drafts.

and individual characteristics such as aspirations and motivations. These two facets of mobility experiences may therefore have differing effects on family formation. In this paper, I present and test a model of the relationship between social mobility and fertility that reconciles some of the contradictory predictions found in earlier literature by acknowledging the dual nature of mobility experiences.

VARIANTS OF THE SOCIAL MOBILITY AND FERTILITY HYPOTHESIS

Five overlapping perspectives provide a rationale for a possible influence of mobility on fertility.[1] The *socialization* perspective views any differences between the fertility behaviors of mobile and nonmobile couples as due only to socialization in two social classes: the class of origin and the class of destination. "Both [origin and destination] groups exert some influence over mobile individuals, since they have, or have had, social contacts with members of both, being placed by economic circumstances amidst the one, while having been socialized among the other. Hence their behavior is expected to be intermediate between that of the two nonmobile classes" (Blau, 1956:291). Since, in this perspective, origin and destination effects suffice to explain the fertility of mobile couples, it is often considered the "null hypothesis" for the following three perspectives.

The *stress* perspective (the first predicting a "mobility" effect) sees the actual mobility experience as being disruptive. Above and beyond the effects of dual socialization, the family size of mobile couples is restricted in response to the stresses and strains experienced in attempting to maintain social membership in two classes rather than one, being marginal to both classes, and/or attempting to gain membership in the class of destination if upwardly mobile or attempting to maintain prior class affiliation if downwardly mobile.

The *isolation* perspective springs from the dissociative hypothesis in stratification research (cf. Ellis and Lane, 1963; Sorokin, 1927). Mobile couples are viewed as socially isolated: they have weakened ties to their class of origin but have difficulty establishing full membership in the class of destination. To compensate for their social isolation, mobile couples increase their family size.

The *status-enhancement perspective*, based on research by Westoff (1953), predicts lowered fertility—at least for upwardly mobile couples—but the rationale differs from the stress perspective. Here both lowered fertility and upward mobility are the consequences of a complex of attitudes and aspirations concerning socioeconomic achievement. Since childbearing and rearing consume resources (e.g., time and money) which aid in the realization of aspirations, couples oriented towards upward mobility will restrict their family size.

The *relative economic* perspective views children as consumer durables supplying parents with psychological and psychic satisfaction, but having heavy economic costs attached (Becker, 1960). Here the general prediction is that individuals with higher economic resources, all other factors being equal, have more children. There are several variants of this particular perspective, differing mainly in the economic standard against which individuals are assumed to compare their current level of economic resources. By assuming that the underlying preferences for children and other material goods are formed in the family of origin, Easterlin's work (1975, 1973) emphasizes the comparison between current or permanent (lifetime) earnings and the level of economic resources available in the parental family. Freedman's (1963) work, on the other hand, by assuming that individuals sharing a relatively homogeneous socioeconomic environment have similar tastes for children and other goods, emphasizes

[1] Fertility has also been hypothesized to influence mobility. The "selection" hypothesis, for instance, suggests that upward mobility is impeded by large family sizes and the probability of downward mobility exacerbated (e.g., Fisher, 1929; Blau and Duncan, 1967). However, numerous economic studies have demonstrated little or no effect of family size on husband's employment (but see Lindert, 1978: Appendix B), so this paper concentrates on the possible influence(s) of socioeconomic mobility on fertility.

the contrast between individuals' economic resources and the average or expected income of an individual in a similar social environment. In the first variant then, the comparison reaches backwards to the family of origin, while in the second, the comparison draws on the economic resources of a current socioeconomic peer group. Although the economic perspective is generally used to assess cross-sectional fertility differences, upward socioeconomic mobility presumably yields a higher level of economic resources relative to the family of origin (in the case of intergenerational mobility) or higher economic resources relative to some socioeconomic peer group. Therefore this perspective suggests that upwardly mobile people should have larger family sizes than otherwise, and downwardly mobile people, smaller family sizes.

In general, the empirical literature offers little evidence for the existence of mobility effects predicted by the second, third, or fourth theoretical perspectives. Analyzing the same data, Berent (1952) and Duncan (1966) conclude that the only consequence of mobility regarding fertility lies in dual socialization, although Hope (1971), reanalyzing the same data but specifying a more exact null hypothesis (the "halfway" hypothesis), disagrees. Studies based on professional occupational groups tend to accept the socialization perspective, although the results often show a negative relationship, but one which is not statistically significant (Boggs, 1957; Perucci, 1967; Scott, 1958; Tien, 1961). Blau and Duncan investigate the relationship between fertility and mobility for all possible combinations of inter- and intragenerational mobility using a national sample encompassing all occupational groups. Although their results imply rejection of the socialization hypothesis, they are reluctant to accept any alternative, since the data do not fall into any consistent pattern (1967: 381). Boyd (1973) suggests that the lack of evidence supporting unique mobility effects may lie in the common nature of mobility experiences in industrialized societies that are characterized by fluid stratification systems. But her results, based on data from five Latin American countries, deem the socialization model adequate. Featherman (1973) tests whether couples who are oriented towards mobility restrict their fertility. His three indices of achievement-related orientations were virtually uncorrelated with childbearing over the study period. Bean and Swicegood's (1979) analysis suggested that downwardly mobile couples limit only the number of unintended births, and not the number of intended.

Switching to the economic perspective, the empirical literature is equivocal. Easterlin has pointed out parallels between the economic positions of cohorts of young men relative to the economic positions of their fathers and lagged fertility rates (1969; 1973; 1975). Studies using individual level data, however, have shown no relationship between family size and economic status relative to family of origin (Olneck and Wolfe, 1978; MacDonald and Rindfuss, 1978). Freedman's (1963) results based on the 1955 Growth of American Families study showed a positive relationship between family size and the contrast between husband's observed and "expected" income: the income predicted by his age, education, and occupation. However, Bernhardt (1972), using Swedish data, found that the relationship was confined to couples at the higher end of the income distribution. She suggested her results differed from Freedman's because her sample was not confined to family planners as was Freedman's sample. Westoff and Ryder (1977:299–302) found little support for Bernhardt's suggestion that only wanted births and relative economic status are positively related: results from the 1970 National Fertility Study showed a negative relationship between husband's relative economic status and both wanted and unwanted fertility.

THE RELATIONSHIP BETWEEN SOCIAL MOBILITY AND FERTILITY: TWO EFFECTS IN ONE

This collection of inconsistent results has led researchers to advise that the search for mobility effects be abandoned (Blau and Duncan, 1967; Boyd, 1973) or at least

suspended until the mobility and fertility hypothesis is better specified regarding the ordering of causal relations and the intervention of factors such as the probability of marriage, timing of marriage, fertility spacing, and the different components of fertility (Bean and Swicegood, 1979; Hargens et al., 1978; Hope, 1971; Perucci, 1967; Scott, 1958; Tien, 1961). Before dismissing the social mobility and fertility hypothesis, however, it seems reasonable to investigate the possibility that the inconsistent results found in previous empirical work resulted from the simultaneous operation of effects proposed by two or more of the theoretical perspectives reviewed above.

Each of the perspectives predicts only one effect of mobility on fertility because each perspective presumes that mobility, as the contrast between an origin and a destination status, is a unitary construct. The total contrast between origin and destination statuses, however, incorporates two different types of mobility experiences, each attributable to a different set of factors. The first component of total mobility, here labeled *mean mobility*, is a function of those factors propelling individuals from certain social origin classes towards a particular social destination. The second component, *relative mobility*, is a function of those factors attributable to the individual's own endeavors, motivations, skill, and/or luck. These two components of total mobility, mean and relative mobility, because they are the product of differing factors, may have differing effects on fertility behavior. The results of previous empirical research, which demonstrate no consistent relationship between fertility and total (unpartitioned) mobility, suggest that if both relative and mean mobility do influence fertility behavior, the effects differ in sign, since differences in sample composition or the sensitivity of differing analytic techniques would then combine these opposite effects in unique ways to produce the inconsistent results found in the empirical literature.

In Table 1, the predictions of the various perspectives are summarized, but here total mobility is partitioned into relative and mean mobility so that two (or more) mobility effects may exist simultaneously. For example, the isolation and stress perspectives could be combined to yield a positive effect of mean mobility and a negative effect of relative mobility on fertility—although there is no easily apparent rationale for not also hypothesizing the opposite set of effects. The combination of the status-enhancement and relative economic perspectives is more appealing. The status-enhancement perspective predicts that an individual's socioeconomic aspirations depress family size through the devotion of economic and personal resources to career advancement rather than to family building. Thus we would expect only that portion of mobility generated by efforts to realize socioeconomic aspirations—relative mobility—to be negatively related to family size. We would also expect this negative effect to be confined to the years when an individual's family building and career building

Table 1. Summary of Mobility Effects on Fertility

	Perspective				
	Socialization up/down	Stress up/down	Isolation up/down	Status enhancement up/down	Relative economic up/down
A. Origins to first occupation					
Mean mobility	0/0	(−/−)	(+/+)	0/0	+/−
Relative mobility	0/0	−/−	+/+	0/0	+/−
B. Origins to current occupation					
Mean mobility	0/0	(−/−)	(+/+)	0/0	+/−
Relative mobility	0/0	−/−	+/+	−/−	+/−
C. First to current occupation					
Mean mobility	0/0	(−/−)	(+/+)	0/0	+/−
Relative mobility	0/0	−/−	+/+	−/−	+/−

NOTE: Parentheses indicate instances in which the perspective does not explicitly predict an effect, but it seems plausible that one might exist.

would be most likely to conflict. On the other hand, if an individual is upwardly mobile without any undue personal effort—that is, his upward mobility is entirely accounted for by mean mobility—there is no reason to expect a limited family size. In fact, the first variant of the relative economic perspective, which emphasizes the contrast between current economic status and economic status of the parental household, predicts that family size will increase with mean upward mobility, since the added economic resources due to mean mobility allow an individual to fully indulge his tastes for children.

Although Table 1 suggests additional combinations of perspectives that also yield contrary predictions of the effects of relative and mean mobility on fertility, the combination of the status-enhancement and relative economic perspectives is the most appealing for two reasons. First, unlike other possible combinations of perspectives, the relative economic and status-enhancement perspectives predict only one set of effects: a positive effect of mean mobility and a negative effect of relative mobility respectively, and not vice versa, because mean mobility—since it involves no extra effort or "striving" on the part of individuals—cannot be interpreted as exerting a negative effect on family size due to status enhancement. Second, this particular combination offers an explanation of the pattern of results found in the previous empirical literature. If, as suggested, fertility is negatively related to relative mobility and positively related to mean mobility, we would expect to find cohort fertility patterns analogous to those found by Easterlin (1973), since the sum of circulation or relative mobility experiences is constrained to equal zero (see below). We would also expect in this case not to find any consistent relationship between fertility and mobility at the individual level in analyses not separating these two components of mobility, since the predicted effects differ in sign.

THE PARTITIONING OF MOBILITY

Occupational mobility can be partitioned into structural (mean) mobility and circulation (relative) mobility. Structural mobility can be thought of as being due to changes in the distributions of the variables used to define mobility. For instance, the shrinking of the agricultural sector implies fewer farmers in one generation than in the previous; therefore, some individuals in the later generation are "forced" to become nonfarmers. Circulation mobility is the mobility remaining after marginal shifts have been taken into account. Generally, stratification research controls for structural mobility, but fertility research does not.[2]

Socioeconomic mobility can be partitioned in a similar manner to occupational mobility. The socioeconomic mobility experiences of an origin group of sons reflect both the general shift in the marginals between the origin and destination statuses of the sons and the sum of the individual circulation mobility experiences:

$$S_i - F = (S_i - \overline{S}) + (\overline{S} - F)$$

---

[2] In fertility research the effects of structural and circulation mobility are generally confounded. For example, a common statistical formulation of the "socialization" model is:

$$Y_{ij} = \overline{Y} + S_i + F_j + e_{ij}$$

where Y_{ij} = the mean fertility in the cell ij,
$\overline{Y}$ = the grand mean,
S_i = the effect of being in the *i*th destination class,
F_j = The effect of being in the *j*th origin class, and
e_{ij} = error or the interaction between the effects of origin and destination classes (cf. Duncan, 1966).

Thus the effect of being in row F_1 and column S_i is a summation of the effects of being in cells 11, 12, 13 . . . 1i. Rearranging terms slightly and still referring to row F_1,

$$\begin{aligned} Y_{1i} &= \overline{Y} + [(S_1 - F_1) + (S_2 - F_1) \ldots (S_i - F_1)] \\ &= \overline{Y} + \Sigma(S_i - F_1) \\ &= \overline{Y} + \Sigma[(S_i - \overline{S}) + (\overline{S} - F_1)] \end{aligned}$$

Thus the effects ascribed to being in row F_1 and column S_i confound the effects of the mean mobility of sons from origin F=1 (the terms in the second set of parentheses) and the circulation mobility of sons from this origin group (the terms in the first set of parentheses).

Bean and Swicegood (1979) attempt to take into account structural mobility due to the socioeconomic upgrading of the labor force over time. Their approach, which consists of using a more stringent criterion for upward mobility than for downward, neglects the fact that structural mobility differs by point of origin (cf. McClendon, 1977).

where S_i = the son's socioeconomic status,
F = the son's father's socioeconomic status, or the son's origin status, and
$\overline{S}$ = the mean socioeconomic attainment of sons from the origin group.

Essentially we are looking at one row of an outflow table, or the distribution of the mobility experiences of one socioeconomic origin group of sons. The term $(S_i - \overline{S})$ refers to the mobility experienced by the ith son relative to the mean attainment of sons from the origin group, while $(\overline{S} - F)$ refers to the mean mobility experience of the origin group of sons.

In this separation of total mobility into its two components, mean and relative mobility, mean or expected destination should reflect factors that are beyond an individual's control, e.g., the relative size of his origin group which may lead to labor-market advantages, or the shrinking of the blue-collar segment of the labor force. The factors influencing an origin group's eventual mean socioeconomic placement are of two types: characteristics of the group such as social background, size, and early socioeconomic achievements; and the group's ability to translate these characteristics into later socioeconomic achievements—an ability which is influenced by "structural" factors such as current labor-market conditions (or, in another terminology, "cohort" and "period" effects). Mean mobility is thus the difference between expected or mean socioeconomic destination and origins, where mean or expected destination reflects the factors above. Relative mobility is the contrast between expected or mean destination and the observed destination. It reflects the variance of socioeconomic achievement within the origin groups. Since the factors influencing the group as a whole have been taken into account, the variance within the origin group's social destinations reflects only individual differences in motivations, skills, and efforts.

DATA AND METHODS

The Occupational Changes in a Generation Study[3] offers an opportunity to test whether or not mobility has dual effects on fertility, since the data set includes measures of origin and several destination statuses as well as family size. The analysis is restricted to white males, aged 35–64, in their first marriage. The age restriction ensures that the respondents have had time to be ensconced in their labor-force career as well as ensuring that their families are close to being completed. The Duncan SEI score for the respondent's father's occupation is used as the measure of the respondent's socioeconomic origins (X), while the SEI scores for the respondent's first and current occupations (Y1 and Y2 respectively) are used as measures of his destination statuses. We could also have considered the respondent's wife's mobility experiences, since they need not be parallel to those of her husband. However, previous research has framed the social mobility–fertility hypotheses with respect to the husband's occupational career. In the effort to suggest that mobility effects can be traced to both structural and individual components, contrary to the implicit assumptions in previous research, and given that many wives are not working, we map only the husband's socioeconomic career.

The two measures of destination status allow the partitioning of total intergenerational mobility from (A) origins to first occupation, and (B) origins to current occupation, as well as the partitioning of total intragenerational mobility, (C) first to current occupation. The differences between the two measures of observed destination and origins ($Y1-X$ and $Y2-X$) are the two measures of total intergenerational mobility. For each of these two forms of intergenerational mobility, mean mobility is the difference between expected or mean destination and origins ($\hat{Y}1-X$; $\hat{Y}2-X$), while relative mobility is the contrast between the respondent's actual destination and expected or mean destination ($Y1-\hat{Y}1$; $Y2-\hat{Y}2$). For both forms of intergenerational mobility, mean or expected destination ($\hat{Y}1$ and $\hat{Y}2$) is defined as a function

[3] See Blau and Duncan (1967) for a full description of the study.

of origin status (X), respondent's education measured in completed years (EDH), five-year birth-cohort membership measured with dummy variables (BC_j), and all interaction terms involving birth-cohort membership (see appendix).

Intragenerational mobility (Y2−Y1) is partitioned in a similar manner to intergenerational mobility, but here mean mobility ($\hat{Y}2-Y1$) is the difference between expected destination and first occupation, while relative mobility is the difference between actual and expected destination ($Y2-\hat{Y}2$). In this case expected destination is a function of the respondent's first occupational socioeconomic status (Y1) as well as of education and all interaction terms involving birth-cohort membership.

This method of partitioning the three measures of total mobility has a strong advantage. In each of the three instances, the covariance between the two components of total mobility, relative mobility and mean mobility, is zero.[4]

Following Duncan's (1966) oft-quoted statement about the need to control for both origin and destination effects before beginning the search for mobility effects, measures of both origin and destination statuses are included in all regression models. It is impossible, however, to estimate separate coefficients for each origin and destination status and the two mobility components; the regression model is, then, underidentified. To overcome this difficulty, the effects of origin and destination statuses on fertility are constrained to be equal, a procedure suggested by Hope (1971). This procedure also has the advantage of specifying a null hypothesis corresponding to the dual socialization perspective in which the fertility of mobile couples is influenced equally by socialization in their class of origin and class of destination.

The first fertility measure used is completed family size. Since there are two transitions in family building of particular interest, the transition from zero to one child and the transition from two to three, the second and third fertility measures used are whether the couple is childless (coded 1 = yes; 0 = no) and whether the couple has three or more children (1 = yes; 0 = no). In the analysis, controls for farm origins of husband and wife (FARMH and FARMW), husband's and wife's years of education (EDH and EDW), duration of marriage (DUR), age of wife (AGEW), husband's age at marriage (AGEMARR), and whether the husband attended a parochial school (PAROCH)—an indicator of Catholicism—are introduced. The means and standard deviations for all variables and their correlations with the three measures of fertility are presented in Table 2.

SUMMARY OF RESULTS

In general, hypotheses about the possible effect of mobility on fertility have not been supported empirically. One way to refine hypotheses about social mobility is to acknowledge the dual character of mobility experiences. Table 3 contrasts the results from the regression of fertility on total mobility and the regression of fertility on the two components of total mobility: mean mobility, that component shared by others with similar background characteristics, and relative mobility, that component unique to an individual.[5] The effect of total mobility on family size is statistically significant at only the .05 level when considering the two forms of intergenerational mobility; when considering intragenerational mobility, the effect is insignificant at any conventional level. When total mobility is partitioned, how-

[4] If Y is a measure of observed destination, X a measure of origins and $\hat{Y}$ is the "expected" or predicted value from the OLS regression of Y on X plus other independent variables (See Appendix) then the covariance of relative mobility ($Y - \hat{Y}$) and mean mobility ($\hat{Y} - X$) is

$$(1) = \mathrm{COV}[(Y - \hat{Y}), (\hat{Y} - X)]$$

$$(2) = \sum_{i=1}^{n} [(Y_i - \hat{Y}_i) - \sum_{i=1}^{n} (Y_i - \hat{Y}_i)/N]$$

$$[(\hat{Y}_i - X_i) - \sum_{i=1}^{n} (\hat{Y}_i - X_i)/N]/N$$

By definition $\Sigma(Y_i - \hat{Y}_i) = 0$ using OLS regression, so expression (2) reduces to zero.

[5] Preliminary analysis showed no statistically significant nonlinearities (at the .001 level) in the effects of relative and mean mobility on fertility, so only linear effects are presented.

Table 2. Means, Standard Deviations, and Correlations with Family Size for All Variables

Variable	Mean	Standard deviation	Correlation with family size	Correlation with childlessness	Correlation with large family
Family size	2.615	1.895	1.000	—	—
Childlessness	.113	.316	−.492	1.000	—
Large family	.458	.498	.760	−.328	1.000
Mobility, origins to first occupation					
Relative	−.046	16.949	−.026	.012	−.033
Mean	−.778	15.195	−.006	−.028	−.006
Total	−.824	22.779	−.024	−.010	−.029
Sum of statuses	54.446	35.419	−.121	−.064	−.074
Mobility, origins to current occupation					
Relative	.089	19.125	−.040	−.009	−.040
Mean	13.049	16.424	−.026	−.031	−.012
Total	13.138	25.232	−.047	−.027	−.038
Sum of statuses	68.409	38.310	−.129	−.019	−.013
Mobility, first to current occupation					
Relative	.017	18.761	−.036	−.007	−.003
Mean	13.855	12.784	.001	−.025	.025
Total	13.962	22.780	−.029	−.019	−.013
Sum of statuses	67.585	39.373	−.139	.040	−.091
Control variables					
FARMH	.327	.469	.104	−.043	.070
FARMW	.295	.456	.093	−.016	.056
EDH	10.883	3.382	−.143	.030	−.068
EDW	11.021	2.820	−.158	.029	−.075
PAROCH	.146	.353	.044	.015	.047
DUR	22.033	8.708	.060	−.034	.000
AGEW	44.342	8.650	−.109	.134	−.134
AGEMARR	25.338	5.194	−.229	.233	−.184

Table 3. Effects of Total, Relative, and Mean Mobility on Children Ever Born

Variables in regression model	Metric coefficients	Partial F-value	R^2	F-value for partitioning of total mobility
A. Origins to first job				
1. Sum of statuses	−.006***	117.53	.015	103.11
Total mobility	−.002*	5.25		
2. Sum of statuses	−.013***	220.75	.028	
Relative mobility	.010***	44.57		
Mean mobility	−.018***	97.99		
B. Origins to current job				
1. Sum of statuses	−.006***	121.91	.017	59.35
Total mobility	−.002*	6.16		
2. Sum of statuses	−.009***	181.29	.025	
Relative mobility	.006***	18.29		
Mean mobility	−.011***	58.59		
C. First to current occupation				
1. Sum of statuses	−.007***	148.47	.019	45.21
Total mobility	−.003	.11		
2. Sum of statuses	−.009***	190.99	.025	
Relative mobility	.006***	19.87		
Mean mobility	−.011***	35.32		

* Coefficient significant at .05 level
*** Coefficient significant at .001 level

Table 4. Net Effects of Relative and Mean Mobility on Children Ever Born

Variables	Origins to first occupation	Origins to current occupation	First to current occupation
Relative mobility	.018***	.021	.014
Mean mobility	−.032***	−.036*	−.028*
Sum of statuses	−.020***	−.023	−.016*
FARMH	.229***	.225***	.201***
FARMW	.153**	.147**	.147**
EDH	.120***	.202	.131
EDW	−.064***	−.063**	−.063***
PAROCH	.365***	.367***	.369***
DUR	.066***	.062***	.059***
AGEW	−.077***	−.076***	−.076***
AGEMARR	−.029***	−.344***	−.036***
Constant	5.555	5.843	6.290
R^2	.113	.112	.112

* Coefficient significant at .05 level
** Coefficient significant at .01 level
*** Coefficient significant at .001 level

ever, the F statistics for the partitioning of mobility are well above the criterion values used to assess statistical significance, and the coefficients for relative and mean mobility for all forms of mobility are significant at the .001 level. Moreover, as predicted, the effects of relative and mean mobility on family size differ in sign: in every case, the coefficient for relative mobility is positive, and the coefficient for mean mobility is negative. These relationships, however, are more prominent when considering intergenerational social movement from origins to first occupation when controlling for other factors known to affect family formation (Table 4). When the dependent variables are the probability of being childless and the probability of having three or more children (rather than family size), the same pattern is observed: the effects of relative and mean mobility are most evident for movement from social origins to first occupation (see Table 5).

DISCUSSION OF RESULTS

Partitioning total mobility substantially aids in the description of the relationship between fertility and social mobility, confirming the impression that there are two components to the relationship between mobility and fertility. Although the amount of variance explained by the mobility components is small, it compares favorably with the explicative power of total mobility found both here and in earlier literature.

Table 5. Net Effects of Relative and Mean Mobility on the Probability of Being Childless and the Probability of Having a Large Family

	Origins to first occupation	Origins to current occupation	First to current occupation
Probability of being childless			
Relative mobility	−.005***	−.005*	−.000
Mean mobility	.006***	.006*	+.000
Sum of statuses	.005***	.005*	+.000
R^2	.092	.089	.087
Probability of having a large family			
Relative mobility	.003*	.006	.003
Mean mobility	−.006**	−.010*	−.006
Sum of statuses	−.004**	−.007*	−.004*
R^2	.073	.072	.071

NOTE: All effects are net of the control variables (see Table 4).
* Coefficient significant at .05 level
** Coefficient significant at .01 level
*** Coefficient significant at .001 level

The direction of the relationships found between family size and the two mobility components are, however, at first glance surprising: fertility is consistently negatively related to mean mobility and positively related to relative mobility. This is the mirror image of the effects hypothesized by the combination of the relative economic and status-enhancement perspectives. Although we had originally drawn on the relative economic perspective to predict a positive association between family size and mean mobility, we can draw upon this perspective to explain the positive effect of relative mobility on fertility. This positive effect can be explained by merging elements of both variants of the relative economic perspective. The first variant emphasizes the contrast between socioeconomic origins and destination, while the second emphasizes the contrast between observed and expected economic resources where expected level of economic resources is generally defined as the average level of resources enjoyed by the husband's current socioeconomic peers. In this analysis, however, expected destination refers to the mean socioeconomic destination of the husband's socioeconomic *origin* peers. Therefore relative mobility taps the contrast between the husband's observed destination and the mean destination of others sharing certain background characteristics. The positive relationship between relative mobility and family size thus implies that if the comparison between the husband's socioeconomic achievement and the mean achievements of his socioeconomic origin group (i.e., those he "started out with") is a favorable one, then these couples are less hampered by economic considerations in their family building.

This particular explanation of the positive relationship between relative mobility and family size also suggests several reasons why the association is most prominent when considering mobility from origins to first occupation. It seems plausible that this particular comparison group, individuals sharing socioeconomic origin characteristics, might be more salient upon initial entry into the labor force than later, since the passage of time would probably weaken the ties and contacts with those sharing only background characteristics. On the other hand, perhaps the contrast between an individual's early socioeconomic achievement and the early achievements of his origin group is more telling than later contrasts because early fertility is such a strong determinant of eventual family size (Bumpass and Mburugu, 1977).

The negative association between mean mobility and fertility is more difficult to explain because none of the perspectives outlined in the literature review offers a compelling rationale. For example, although the stress hypothesis predicts a negative relationship between mean mobility and fertility, it does not preclude expecting a similar relationship between relative mobility and fertility, because this perspective assumes that any mobility experience is stressful. However, it has been suggested that upward mobility from class of origin is associated with "over-conformity" to the attitudes and behaviors found in the class of destination (e.g., Kessin, 1971; Tumin, 1967). On the whole, the respondents in this sample benefited from the socioeconomic upgrading of the labor force between the time of their father's generation and their own ($\overline{Y2} - \overline{X} = 12.95$). There is also a negative relationship between fertility and socioeconomic status. Perhaps the respondents adopted the attitudes and behaviors of their destination classes more than would be expected given the dual socialization hypothesis. That is, rather than being equally responsive to the attitudes and behaviors of both their origin and destination classes, they chose to conform more closely to those of their class of destination. Alternatively, since contraceptive use varies directly with socioeconomic status (Whelpton et al., 1966), the dampening effect of mean mobility on family size could also be a function of the general upward movement into socioeconomic environments in which family planning is more widely known and practiced, allowing couples to limit the number of unintended births. This argument is supported by Bean and Swicegood's (1979) results showing that

only unintended—and not intended—births were negatively related to (total) mobility and Kantner and Kiser's results showing a "greater regularity of contraception among upwardly mobile couples" (1954:103).

CONCLUSIONS

Partitioning total mobility into relative mobility, the portion attributable to an individual's unique characteristics, and mean mobility, the portion shared by others with similar socioeconomic origin characteristics, shows that fertility behavior is negatively related to mean mobility and positively related to relative mobility.

The relative economic perspective explains the positive effect of relative mobility on family size by assuming that an individual's fertility behavior is responsive to his socioeconomic achievement relative to the achievements of those persons sharing his background characteristics—that added or extra economic resources enable him to indulge more fully his tastes for children. To explain the negative effect of mean mobility on family size, I suggest that an individual's fertility behavior is more sensitive to the attitudes and/or behaviors in his destination class than his origin class. Since the contrasts between origins, early mean or expected socioeconomic achievement, and early actual socioeconomic achievement are the most telling, I then suggest that the effects of the contrasts between an individual's achievement and the average achievements of his origin group recede with the passage of time, or that early mobility experiences strongly affect the onset and subsequent pacing of fertility, which then in turn strongly affect completed family size.

The evidence presented in this paper for the existence of two simultaneous and contrary effects of mobility on fertility suggests several advantages in the partitioning of mobility. First, we may find that other attitudes and behaviors outside the realm of family building are sensitive to more than the simple contrast between two social statuses—that previous research shows only scanty evidence for mobility (and status inconsistency) effects because contrary effects of mean and relative mobility are cancelling one another. Second, the partitioning of mobility may allow easier comparisons of, for example, relative mobility effects across time and space, since mean mobility, which is largely generated by contextual factors, has been taken into account.

APPENDIX

For both types of intergenerational socioeconomic mobility (origins to first occupation, and origins to current occupation), mean mobility is defined as the difference between expected destination and origins ($\hat{Y}_1 - X$; $\hat{Y}_2 - X$), while relative mobility is the difference between expected destination and observed destination ($Y1 - \hat{Y}1$; $Y2 - \hat{Y}2$). The respondent's expected destination scores are the predicted values from the regression of first (FJSES) and current occupation (CJSES) status scores on the respondent's father's occupational status score (PSES), the respondent's education (EDH), and all interaction terms involving birth cohort membership (BC_i):

$$
\begin{aligned}
\hat{Y}1 = \widehat{FJSES} = {} & .2071\,(BC_1 \times PSES) \\
& + .2136\,(BC_2 \times PSES) \\
& + .2159\,(BC_3 \times PSES) \\
& + .2155\,(BC_4 \times PSES) \\
& + .2257\,(BC_5 \times PSES) \\
& + .2466\,(BC_6 \times PSES) \\
& + 2.8106\,(BC_1 \times EDH) \\
& + 2.5838\,(BC_2 \times EDH) \\
& + 2.6208\,(BC_3 \times EDH) \\
& + 2.9447\,(BC_4 \times EDH) \\
& + 3.0880\,(BC_5 \times EDH) \\
& + 2.9725\,(BC_6 \times EDH) \\
& - 9.5658 \qquad R^2 = .343
\end{aligned}
$$

$$
\begin{aligned}
\hat{Y}2 = \widehat{CJSES} = {} & .1894\,(BC_1 \times PSES) \\
& + .1631\,(BC_2 \times PSES) \\
& + .1731\,(BC_3 \times PSES) \\
& + .1940\,(BC_4 \times PSES) \\
& + .2394\,(BC_5 \times PSES) \\
& + .2320\,(BC_6 \times PSES) \\
& + 3.7727\,(BC_1 \times EDH) \\
& + 3.9006\,(BC_2 \times EDH) \\
& + 3.8335\,(BC_3 \times EDH) \\
& + 3.9019\,(BC_4 \times EDH) \\
& + 3.8239\,(BC_5 \times EDH) \\
& + 3.8893\,(BC_6 \times EDH) \\
& + 6.5052 \qquad R^2 = .386
\end{aligned}
$$

For intragenerational mobility (first to current occupation) mean mobility is defined as the difference between expected destination and first occupation, while relative intragenerational mobility is the contrast between expected and observed destination. In this case, expected destination is:

$$
\begin{aligned}
\hat{Y}2 = \text{CJSES} = {} & .2970\ (BC_1 \times FJSES) \\
& + .2609\ (BC_2 \times FJSES) \\
& + .2738\ (BC_3 \times FJSES) \\
& + .3167\ (BC_4 \times FJSES) \\
& + .2914 (BC_5 \times FJSES) \\
& + .4440\ (BC_6 \times FJSES) \\
& + 3.2830\ (BC_1 \times EDH) \\
& + 3.4825\ (BC_2 \times EDH) \\
& + 3.3984\ (BC_3 \times EDH) \\
& + 3.3303\ (BC_4 \times EDH) \\
& + 3.3961\ (BC_5 \times EDH) \\
& + 3.0445\ (BC_6 \times EDH) \\
& - 3.8779 \qquad R^2 = .410
\end{aligned}
$$

REFERENCES

Bean, Frank and Gray Swicegood
1979 "Intergenerational occupational mobility and fertility: a reassessment." American Sociological Review 44:608–19.

Becker, Gary
1960 "An economic analysis of fertility." Pp. 209–31 in Universities National Bureau Committee for Economic Research, Demographic and Economic Change in Developed Countries. Princeton University Press: Princeton.

Becker, Gary and Nigel Tomes
1976 "Child endowment and the quantity and quality of children." Journal of Political Economy 84:S143–62.

Berent, Jerzy
1952 "Fertility and social mobility." Population Studies 5:244–60.

Bernhaardt, Eva
1972 "Fertility and economic status: some recent findings on differentials in Sweden." Population Studies 26:175–84.

Blalock, Hubert M.
1967 "Status inconsistency, social mobility, status integration and structural effects." American Sociological Review 32:790–801.

Blau, Peter M.
1956 "Social mobility and interpersonal relations." American Sociological Review 21: 290–5.

Blau, Peter and Otis Dudley Duncan
1967 The American Occupational Structure. John Wiley: New York.

Boggs, S.
1957 "Family size and social mobility in a California suburb." Eugenics Quarterly 4:208–13.

Boyd, Monica
1973 "Occupational mobility and fertility in metropolitan Latin America." Demography 10:1–17.

Bumpass, Larry and Edward Mburguru
1977 "Age at marriage and completed family size." Social Biology 24:31–7.

Duncan, Otis Dudley
1966 "Methodological issues in the analysis of social mobility." Pp. 51–97 in Neil Smelser and Seymour Lipset (eds.), Social Structure and Mobility in Economic Development. Chicago: Aldine.

Easterlin, Richard A.
1975 "An economic framework for fertility analysis." Studies in Family Planning 6:54–63.
1973 "Relative economic status and the American fertility swing." Pp. 170–235 in E. B. Sheldon (ed.), Family Economic Behavior. Philadelphia: Lippincott.
1969 "Towards a socioeconomic theory of fertility: survey of recent research on economic factors in American fertility." Pp. 127–56 in B. J. Behrman, Leslie Corsa, and Ronald Freedman (eds.), Fertility and Family Planning: A World View. Ann Arbor: University of Michigan Press.

Ellis, Robert A. and W. Clayton Lane
1963 "Structural supports for upward mobility." American Sociological Review 28:743–56.

Featherman, David L.
1973 "Marital fertility and the process of socioeconomic attainment: an examination of the mobility hypothesis." Pp. 104–31 in Larry Bumpass and Charles Westoff (eds.), The Later Years of Childbearing. Princeton: Princeton University Press.

Fisher, R. A.
1929 The Genetic Theory of Natural Selection. New York: Dover.

Freedman, Deborah
1963 "The relation of economic status to fertility." American Economic Review 53:414–26.

Hargens, Lowell L., James C. McCann and Barbara F. Reskin
1978 "Productivity and reproductivity: fertility and professional achievement among research scientists." Social Forces 57:154–63.

Hope, Keith
1971 "Social mobility and fertility." American Sociological Review 36:1010–29.

Kanter, J. F. and C. V. Kiser
1954 "The interrelation of fertility, fertility planning and intergenerational social mobility." Milbank Memorial Quarterly 29:69–103.

Kessin, Kenneth
1971 "Social and psychological consequences of intergenerational occupational mobility." American Journal of Sociology 77:1–18.

Lindert, Peter
1978 Fertility and Scarcity in America. Princeton: Princeton University Press.

McClendon, McKee
1977 "Structural and exchange components of vertical mobility." American Sociological Review 42:56–74.

McDonald, Maurice M. and Ronald R. Rindfuss
1978 "Relative economic status and fertility: evidence from a cross-section." Pp. 291–307 in Julian Simon (ed.), Research in Population Economics. Volume 1. Greenwich: JAI Press.

Olnick, Michael and Barbara Wolfe
1978 "A note on some evidence on the Easterlin Hypothesis." Journal of Political Economy 86:953–8.

Perucci, Carolyn C.
1967 "Social origins, mobility patterns and fertil-

ity." American Sociological Review 32:615–25.

Schultz, Theodore C.
1973 "The value of children: an economic perspective." Journal of Political Economy 81:S2–13.

Scott, Wolf
1958 "Fertility and social mobility among teachers." Population Studies 11:25–61.

Sorokin, Pitirim A.
1927 Social and Cultural Mobility. New York: Free Press.

Tien, H. Yuan
1961 "The social mobility/fertility hypothesis reconsidered: an empirical study." American Sociological Review 26:247–57.

Tumin, Melvin
1967 Social Stratification: The Forms and Functions of Inequality. Englewood Cliffs, NJ: Prentice-Hall.

Westoff, Charles F.
1953 "The changing focus of differential fertility research: the social mobility hypothesis." Milbank Memorial Fund Quarterly 31:24–38.

Westoff, Charles F. and Norman B. Ryder
1977 The Contraceptive Revolution. Princeton: Princeton University Press.

Whelpton, Pascal K., Arthur A. Campbell, and John F. Patterson
1966 Fertility and Family Planning in the United States. Princeton: Princeton University Press.

DYNAMICS OF ORGANIZATIONAL EXPANSION IN NATIONAL SYSTEMS OF EDUCATION*

GLENN R. CARROLL
Brown University and *Zentrum für Umfragen, Methoden und Analysen*

National systems of education expand organizationally at unequal speeds in different countries and time periods. A nonlinear differential equation model based on ecological theory is advanced to explain these differences. The model is tested empirically, including an analysis of local stability. The results suggest that organizational expansion in large, geographically decentralized systems follows a process of logistic growth with a ceiling set by environmental resources.

INTRODUCTION

Expansion in national systems of education has been analyzed traditionally by modernization theorists (e.g., Lipset, 1960). These analysts typically regard education as a primary mechanism by which societies move from premodern to modern social and political structures.

Educational expansion, however, is also an organizational process. As such it entails the mobilization of environmental resources and their assemblage into differentiated structures of formal organization. Too often this fact has been ignored by the modernization theorists, who often implicitly overlook the constraints that impede the rapid expansion of massive educational organizations.

Cross-national data on educational enrollments show that expansion occurs with unequal speed in the countries of the world (see Meyer and Hannan, 1979). Current organizational theory, by demonstrating how organizations are constrained by external conditions, suggests that much of this variation can be explained by national differences in the environments of educational organizations.

The study of educational expansion also contains advantages for the testing of theories about organizations and environments; educational organizations are large-scale hierarchical systems that remain comparable despite cultural dif-

* Direct all correspondence to Glenn R. Carroll, Department of Sociology, Brown University, Providence, RI 02912.

This research was supported in part by NSF Grant SOC78-12315 to Michael T. Hannan and Nancy Brandon Tuma and NIMH Grant 2T32MH15149-03 to the Organizations Research Training Program at Stanford University. I wish to thank Michael T. Hannan, François Nielsen, John Freeman, Barbara Warsavage, Peter Oberto, and the journal referees for helpful comments and advice.

ferences in their environments. Much of the previous research on organizations and environments focuses on smaller organizations with culturally similar environments and ignores interorganizational relations.

I consider here the problem of organizational expansion in national systems of education as an ecological phenomenon. I use the population ecology perspective because it emphasizes resource constraints, interorganizational competition, and temporal disequilibrium. My model follows closely a previous study by Nielsen and Hannan (1977); however, I pay more attention to the time path of expansion processes, the exact form of resource constraints, and the character of interorganizational relations.

AN ECOLOGICAL MODEL OF ORGANIZATIONAL EXPANSION

The population ecology perspective has been detailed by Hannan and Freeman (1977) and Aldrich (1979). I use here the general structure outlined by Nielsen and Hannan (hereafter NH) and repeat only the basic elements of the model.

In NH's characterization, national systems of education are composed of three general organizational *forms:* primary, secondary, and tertiary educational organizations. These organizational forms exhibit increasing work complexity and per-unit investment as the hierarchical ladder from primary to tertiary is ascended. The collection of all organizational units within a form is defined as a *population* of organizations, and the three populations considered jointly constitute an organizational *community*.

Organizational Carrying Capacity

Organizational ecology assumes that organizational expansion is finitely constrained by available environmental resources: that is, for any given level of resources there exists a ceiling, or carrying capacity, on organizational size (Stinchcombe, 1965; Etzioni, 1968; Glassow, 1978). An organization's carrying capacity does not necessarily coincide with its realized size. Instead, it indicates the upper limit on size regardless of the speed at which the organization expands. Thus characterized, carrying capacity must be modelled as an unobservable variable.

For national systems of education, carrying capacity is limited by the level of the available resources of national population and national wealth. Enrollments cannot exceed the population size or an amount sustainable by the level of wealth. For simplicity, NH assumed a linear relationship between the population size of an organizational form and environmental resources. I argue that the relationship should be nonlinear, because, although resources set a ceiling for organizational size, the effect of a unit increment in the level of resources is likely to be more pronounced in resource-scarce environments. Similarly, resource-abundant environments should be less sensitive to small fluctuations in the resource base. These ideas suggest

> Proposition 1: Resource abundance increases organizational carrying capacities at decreasing rates.

The argument can be specified in nonlinear equations:

$$1/\mathrm{PRI}^* = c_{10} + c_{11}(1/R) + c_{12}(1/R^2) \tag{1a}$$

$$1/\mathrm{SEC}^* = c_{20} + c_{21}(1/R) + c_{22}(1/R^2) \tag{1b}$$

$$1/\mathrm{TER}^* = c_{30} + c_{31}(1/R) + c_{32}(1/R^2) \tag{1c}$$

where PRI*, SEC*, and TER* represent the enrollment carrying capacities for primary, secondary, and tertiary organizations, and R indicates any of the resource variables of national population (POP), national wealth (GNP), or primary school age population (PRIPOP, used only for PRI*).[1] Although other functional forms might also specify the argument, I use the forms of Equation 1 because they are computationally convenient.

[1] Like the NH model, this specification predicts non-zero carrying capacities for higher level systems when lower level enrollments are zero. Unlike the NH model, however, the model proposed here does not imply that realized enrollments will also be nonzero.

Relations between Organizational Forms

Interorganizational theorists usually describe the links between organizations according to their overt manifestations (e.g. Thompson and McEwen, 1958). Ecologists, however, characterize the relations between social units with respect to their interdependent viability, regardless of visible behavior. Hawley (1950) concentrates on the development of symbiotic or mutually dependent ties between social units. Hannan and Freeman (1977:940) argue that "as long as the resources which sustain organizations are finite and populations have unlimited capacity to expand, competition must ensue."

Relations between forms of educational organization can thus be one of two general types—symbiotic or competitive—depending on how the viability of one form affects the viability of another. Although lower level educational expansion enhances later expansion of the higher level schools, this symbiotic relation does not preclude competition.[2] Educational organizations compete for many of the same resources of money, personnel, and legitimacy. Because when one organizational form expands, it decreases the resources available to the others, a model of interorganizational forms should allow all possible types of relations. This requires a specification of the form

$$dPRI/dt = f_1(PRI, SEC, TER) \quad (2a)$$
$$dSEC/dt = f_2(PRI, SEC, TER) \quad (2b)$$
$$dTER/dt = f_3(PRI, SEC, TER) \quad (2c)$$

where PRI is the level of primary enrollments, SEC is secondary enrollments, TER is tertiary, and d/dt is the rate of change in a variable over time.

Completely specified models of the form in Equations 2 associate a parameter with each right-hand side variable. When a parameter is positive, it indicates a symbiotic relation with the dependent variable; when negative, a competitive relation. Since we expect both types of effects between forms of educational organizations, predictions of this sort are not feasible. We can, however, examine the total 3×3 matrix of parameters which summarize the community structure between organizational forms. In general, the ties that bind communities are expected to be stronger in periods of adversity (see Pianka, 1978). Mutually dependent social units will exhibit strengthened ties to overcome adversity, and competitive units will have enhanced opportunities to exclude the opponent. In this analysis, adversity is defined as an environment with scarce resources, and we expect

> Proposition 2: The community structure of relations between organizational forms will be more intense in environments with scarce resources (Hannan and Freeman, 1978).

The relational intensity of a community can be measured by calculating the average absolute value of parameters in the endogenous matrix or by directly comparing symmetric coefficients in separate matrices.

Time Trajectory of the Expansion Process in a Large, Geographically Decentralized Organizational System

Organizations adjust to increases in their carrying capacities with unequal speeds. NH argue that speed of adjustment varies with both work complexity and per-unit cost. Complex and differentiated structures require the assemblage of specialized skills and material infrastructure. Mobilization of these items requires more time, coordination, and planning than for simpler, less specialized equipment and personnel. Similarly, for a given input of money, less costly organizations can expand more than others. Thus

> Proposition 3: Both a high degree of complexity and a high cost per unit lower the speed of adjustment of organizations (Nielsen and Hannan, 1977:484).

Since the forms of educational organizations vary monotonically along these dimensions, this proposition leads to the empirical prediction that the adjustment of primary organizational populations will be fastest, followed by secondary and then tertiary. In addition, as NH also argue, adjustment is expected to involve less

[2] NH (1977:490) acknowledge this point but fail to introduce it in their model.

friction and obtain more quickly in resource-abundant environments. Thus

> Proposition 4: Similar types of organizational forms adjust fastest to changes in carrying capacity in resource-abundant environments.

NH's analysis yields general properties of organizational adjustment speeds irrespective of an organization's size. But organizational growth rates may also exhibit density-dependence (see Brittain and Freeman, 1980). In particular, I argue that rapid organizational expansion is impeded in the early periods of system development. Prior to large-scale organizational expansion at the societal level, obstacles that constrain massive resource mobilization must be eliminated. These obstacles are frequently political and require the concentrated efforts of the mobilizing agents. For the development of national education systems, these obstacles include institutional legitimation, resource allocation, and implementation constraints.

Institutional legitimation. A prerequisite of high educational enrollments is the formal acceptance of the institution of education by the major societal actor, usually the state. But a high societal participation rate also frequently requires similar sanctioning by other, often conflicting, societal actors. When, as in most societies, these units are highly dispersed and decentralized, widespread legitimation requires a significant time period. Further, extensive public acceptance may require even more time, especially if communication facilities are underdeveloped. In any case, institutional acceptance requires a redefinition of societal norms, and the process by which this occurs is usually slow and distributed over a wide range of disparate social actors (see Meyer, 1977).

Resource allocation. Even after the goals of an organizational system are institutionalized, competition for developmental resources may delay initial rapid growth. At the early stage of system development, an organization is likely to have a disadvantaged competitive position in comparison to established organizations. With sparse assets and few constituents, the small organizational system is unlikely to muster the political strength necessary for successful claims upon scarce resources, whatever its level of institutional acceptance.[3] As the organizational system attains resources, however, it gains constituents and improves its competitive position.

Implementation constraints. Considerable time is required to mobilize the resources needed in program implementation once the allocation effort has succeeded. In education, these resources include the construction and acquisition of classroom facilities, the development of instructional technology, the recruitment and training of personnel, and the recruitment and retention of students. Pressman and Wildavsky (1973) show that the delays involved can be enormous. These delays are likely to be most pronounced in geographically decentralized systems where expansion emanates from core regions and resembles a diffusion process (see Rogers and Shoemaker, 1971). Given these facts, rapid early organizational expansion seems improbable, even if institutional acceptance and resource allocation are accomplished quickly.

The initial growth process that I have described is constrained by a variety of environmental obstacles. Overcoming these obstacles is most difficult the first time they are encountered. As the organizational system grows and develops experience, these factors decreasingly inhibit expansion. In industry, this phenomenon has been labelled "the learning curve," and its occurrence is well-documented in labor-intensive enterprises (see Baloff, 1971). In education, several empirical studies also point to slow early organizational growth. The adoption of compulsory education laws by individual states of the United States has been shown to be a logistic diffusion process, thereby implying a slow early rate of institutional legitimation (Pemberton, 1936). With respect to

---

[3] A current example of this situation is the national public health-insurance program in the United States. While the goals of this system have been widely accepted for many years, the mobilizers have been unable to attain the resources required for development.

enrollment expansion, Hamblin et al. (1973) show that the number of higher education degrees in the United States have different growth rates in the earlier historical period, and Meyer et al. (1977) show that cross-national enrollment expansion closely resembles a logistic diffusion process. This evidence and my arguments above suggest

> Proposition 5: Geographically decentralized organizational systems expand slowly in early stages of development.

My argument concentrates on impediments during the initial development of the organizational population. NH's discussion focuses on constraints that are most salient near the carrying capacity. When these arguments are coupled, they suggest growth will be least constrained and most rapid at the point equidistant from the constrained regions; namely, the midpoint. Since the logistic growth model has this property, I follow the suggestions of Hannan and Freeman (1977) and Aldrich (1979) and propose that we study the population ecology of educational organizations with this model. For the case when populations also interact, this yields the Lotka-Volterra specification

$$F_1 = (1/PRI)\,(dPRI/dt) = a_1 + b_{11}PRI(t) + b_{12}SEC(t) + b_{13}TER(t) \quad (3a)$$

$$F_2 = (1/SEC)(dSEC/dt) = a_2 + b_{21}PRI(t) + b_{22}SEC(t) + b_{23}TER(t) \quad (3b)$$

$$F_3 = (1/TER)(dTER/dt) = a_3 + b_{31}PRI(t) + b_{32}SEC(t) + b_{33}TER(t) \quad (3c)$$

where a_j is the adjustment speed parameter, b_{jj} are the feedback or growth-dampening parameters, and the b_{jk} coefficients are the population interaction parameters indicating symbiosis and competition.

The empirical predictions are that the adjustment speed parameters will decrease monotonically from PRI to TER, or $a_1 > a_2 > a_3$. We also expect that the subsamples of rich environments will have higher a_1, a_2, and a_3 values than the poor environments. Similarly, we expect that relational intensity$(1/JK \sum_{j=1}^{J} \sum_{k=1}^{K} |b_{jk}|)$ will be greater in poor than in rich environments.

Comparison with the Nielsen and Hannan Model

The model proposed here diverges from the NH model in several important ways. In the formulation of the carrying capacity equation, my model allows nonlinear resource effects, and NH's does not. My model also allows all possible relations to occur between organizational forms, and NH's does not. On both of these points, the model I have specified is more general.

Our models also predict different time paths for the growth process. On this issue, however, I argue that my model, in addition to being motivated on distinct substantive grounds, is actually a better specification of NH's original theoretical argument. To see this, we must recall that a fundamental part of NH's argument relies on per-unit comparisons across organizational forms (see Proposition 3, above).[4] They argue persuasively that per-unit growth adjustments and per-unit resource investments will vary systematically by educational level. Yet the mathematical model which NH propose is not a per-unit growth model; it is a constant-rate growth model. Given their theoretical reliance upon per-unit growth arguments, I argue that the functional form of the dynamic model is misspecified.

NH's mathematical model is typically referred to as the linear partial adjustment model (see Hannan and Tuma, forthcoming). Introducing per-unit growth into this model primarily affects the adjustment speed parameters. Specifically, the adjustment speed parameter becomes multiplied by the enrollment level in a per-unit specification. This shift converts the NH model into a logistic growth model

[4] Also consider the proposition not directly addressed here: "The more complex the core technology of an organization, and the more efforts and resources invested in *each unit* of input, the more selective the boundary-spanning components in charge of input" (Nielsen and Hannan, 1977:485), emphasis mine.

(the terms 1/PRI, 1/SEC, and 1/TER in Equations 3 show the per-unit growth property of the logistic). The logistic and linear partial adjustment models converge in the region near the carrying capacity but differ considerably in their earlier time trajectories. In particular, logistic growth generates an S-shaped population curve which attains maximum growth at half the level of the carrying capacity. The linear partial adjustment model has a convex population curve and grows most rapidly at low population levels, especially the origin. Thus, the major difference between the two models pertains to the rate of growth at the initial stages of the process.

ESTIMATION

The Lotka-Volterra model seems to be a simple extension from the NH linear partial adjustment model, but the implications of its use have serious consequences. The most serious complication arises in estimation; the exact integral solution for the system of Equations 3 has not been found, although it is known to exist. This creates a dilemma. If we choose a linear differential equation model with an explicit solution, then for theoretical reasons we believe the model to be misspecified. If, however, we choose the nonlinear dynamic model in Equation 3, then we are true to theory but lack an explicit solution that can be used to estimate parameters. Since it makes sense to work only with the correct model, I have chosen to estimate the parameters in Equation 3 with an approximate rather than exact solution.

The estimation strategy that I propose involves the use of an "exact discrete approximation." This strategy is rare in sociology but common in disciplines where empirical applications of differential equations are widespread. The strategy consists of first writing a numerically tractable, discrete time approximation to the integral solution of the model. We next determine a set of transformation rules which relate the parameters of the discrete equation to the differential equation. If the limiting behavior of the discrete model converges to the differential equation as the time interval becomes infinitesimally small, then our approximation is "exact."

The exact discrete approximation that I use is due to Leslie (1958), although my exposition follows closely Pielou (1977). For convenience, I show the details of this procedure only for the PRI equation (the others are similar). In this case the discrete time model is

$$\mathrm{PRI}(t+h) = \lambda_p^h \,\mathrm{PRI}(t)/(1 + \gamma_{11}\,\mathrm{PRI}(t) + \gamma_{12}\,\mathrm{SEC}(t) + \gamma_{13}\,\mathrm{TER}(t)) \tag{4}$$

where $\gamma_{ij} = d_{ij}\,(\lambda_p^h - 1)$ and h is defined as the unit for the time t. It now follows that

$$\begin{aligned}(\mathrm{PRI}(t+h) - \mathrm{PRI}(t))/h &= (\mathrm{PRI}(t)(\lambda_p^h - 1)/h) \\ &\quad [(1 - d_{11}\mathrm{PRI}(t) \\ &\quad - d_{12}\mathrm{SEC}(t) \\ &\quad - d_{13}\mathrm{TER}(t))/ \\ &\quad (1 + d_{11}(\lambda_p^h - 1)\,\mathrm{PRI}(t) \\ &\quad + d_{12}(\lambda_p^h - 1)\,\mathrm{SEC}(t) \\ &\quad + d_{13}(\lambda_p^h - 1)\,\mathrm{TER}(t))]\end{aligned} \tag{5}$$

and
$$\lim_{h\to 0}(\mathrm{PRI}(t+h) - \mathrm{PRI}(t))/h = d\mathrm{PRI}/dt \tag{6a}$$

$$\lim_{h\to 0}(\lambda_p^h - 1)/h = \lim_{h\to 0}\lambda_p^h \ln\lambda_p/1 = \ln\lambda_p \tag{6b}$$

by l'Hôspital's rule. Also as $h \to 0$ the denominator in square brackets in Equation 5 becomes 1 so that

$$d\mathrm{PRI}/dt = \mathrm{PRI}(t)\ln\lambda_p\,[1 - d_{11}\,\mathrm{PRI}(t) - d_{12}\,\mathrm{SEC}(t) - d_{13}\,\mathrm{TER}(t)] \tag{7}$$

which is equivalent to Equation 3a when

$$a_1 = \ln\lambda_p \tag{8a}$$
$$b_{11} = -d_{11}\ln\lambda_p \tag{8b}$$
$$b_{12} = -d_{12}\ln\lambda_p \tag{8c}$$
$$b_{13} = -d_{13}\ln\lambda_p \tag{8d}$$

I next extend this model to allow the carrying capacity to depend on exogenous variables. I begin by defining PRI* as the carrying capacity of the PRI equation when the competitors SEC and TER are not present. Using the fact that $\mathrm{PRI}^* = -a_1/b_{11}$ and that $\gamma_{11} = (-b_{11}/a_1)(\lambda_p^h - 1)$, it is easily seen that $\gamma_{11} = (1/\mathrm{PRI}^*)(\lambda_p^h - 1)$. We can thus define

$$(\mathrm{PRI}^*)^{-1} = c_{10} + c_{11}\mathrm{GNP}^{-1} + c_{12}\mathrm{POP}^{-1} + c_{13}\mathrm{PRIPOP}^{-1} + c_{14}\mathrm{GNP}^{-2} \tag{9}$$

and substitute into the γ_{11} term in Equation 4. Now when both sides of the equation are inverted[5]

$$\begin{aligned}\mathrm{PRI}(t+h)^{-1} = {} & c_{10}\lambda_p^{-h}(\lambda_p^h - 1) \\ & + \lambda_p^{-h}\mathrm{PRI}(t)^{-1} \\ & + \lambda_p^{-h}\gamma_{12}\mathrm{PRI}(t)^{-1}\mathrm{SEC}(t) \\ & + \lambda_p^{-h}\gamma_{13}\mathrm{PRI}(t)^{-1}\mathrm{TER}(t) \\ & + c_{11}\lambda_p^{-h}(\lambda_p^h - 1)\mathrm{GNP}^{-1} \\ & + c_{12}\lambda_p^{-h}(\lambda_p^h - 1)\mathrm{POP}^{-1} \\ & + c_{13}\lambda_p^{-h}(\lambda_p^h - 1)\mathrm{PRIPOP}^{-1} \\ & + c_{14}\lambda_p^{-h}(\lambda_p^h - 1)\mathrm{GNP}^{-2}\end{aligned} \tag{10a}$$

or,

$$\begin{aligned}\mathrm{PRI}(t+h)^{-1} = {} & a_0^* + a_1^*\mathrm{PRI}(t)^{-1} \\ & + a_2^*\mathrm{PRI}(t)^{-1}\mathrm{SEC}(t) \\ & + a_3^*\mathrm{PRI}(t)^{-1}\mathrm{TER}(t) \\ & + a_4^*\mathrm{GNP}^{-1} \\ & + a_5^*\mathrm{POP}^{-1} \\ & + a_6^*\mathrm{PRIPOP}^{-1} \\ & + a_7^*\mathrm{GNP}^{-2}\end{aligned} \tag{10b}$$

which is in a form amenable to regression analysis.

Ayala, Gilpin, and Ehrenfeld (1973) have estimated similar models with ordinary least squares, but in the present context probable complications of autocorrelation and heteroscedasticity make this route inappropriate. However, both problems can be corrected by using the weighted generalized least squares estimator (WGLS) applied to pooled time-series of cross-sectional data (see Hannan and Young, 1977; Nielsen and Hannan, 1977; Nielsen, 1978). This estimator is an extension of the random effects estimator of Henderson (1952, 1963) that controls for autocorrelation and heteroscedasticity.[6]

I have applied these techniques to data similar to NH's; our sources are the same, and we both use data for the period 1950–1970. National enrollment levels were obtained from UNESCO (1971) and are measured in hundred thousands. Population figures were taken from Taylor and Hudson (1971) and are given in millions. Gross national product is from IBRD (1971) and is measured in billions of standardized 1964 United States dollars. Rich and poor nation subsamples were determined by separating the sample on median GNP per capita measured in 1950. In contrast to NH, who use only independent nations, I have used all those for which data are available. The logic of the theoretical arguments does not seem to me to require a separation of independent and nonindependent countries. For those who disagree, consolation may be found in the small difference in the sample sizes.

EMPIRICAL FINDINGS

The estimation strategy that I have adopted here is indirect; the first set of parameters must be transformed before the coefficients of interest are found. Since these first parameters have no direct substantive interpretation, I do not discuss them. Instead, I present these estimates in Table 1 for those who wish to follow my calculations, but turn immediately to the coefficients of substantive interest, which are presented in Table 2.[7]

The estimates in Table 2 were calculated by transforming the regression coefficients in Table 1 with the algorithms in Equations 8. The estimates for the speed

---

[5] At this point the computational convenience of the carrying capacity equation (Equations 1) becomes obvious; many other functional forms will not yield a linear equation when inverted. In such cases, an alternative nonlinear estimation technique must be used.

[6] The estimator controls for autocorrelation due to a time-invariant, country-specific component in the disturbance. For heteroscedasticity, following NH, I use PRIPOP as the weighting factor in the PRI equation and POP in the SEC and TER equations. Inspection of the residuals showed these correcting factors to be adequate.

[7] Following current practice in sociological research (see, for example, Hannan and Freeman, 1978; Nielsen, 1980; and Rosenfeld, 1980), I solved for the differential equation coefficients of all variables in the model, even when the indirect estimates from the integral equations have relatively large standard errors. Given the lack of a theoretically based alternative, this practice seems at least temporarily acceptable since the standard errors reported in Table 1 are not estimates for the coefficients of the model under substantive consideration, but instead are estimates for the set of transformed coefficients in an equation which is nonlinear in the parameters and the variables. In later work, I intend to address the issue of standard errors in differential equation models directly.

Table 1. Weighted Generalized Least Squares Estimates of Coefficients in Equation 10b

	Dependent variable					
	$PRI(t+5)^{-1}$		$SEC(t+5)^{-1}$		$TER(t+5)^{-1}$	
	Rich	Poor	Rich	Poor	Rich	Poor
Independent variables						
$PRI(t)^{-1}$	.736	.304				
	(.058)	(.024)				
$SEC(t)^{-1}$			.430	.562		
			(.029)	(.022)		
$TER(t)^{-1}$					.602	.762
					(.022)	(.046)
$SEC(t)PRI(t)^{-1}$	.0006	−.016				
	(.004)	(.084)				
$TER(t)PRI(t)^{-1}$	−.0004	.085				
	(.017)	(.518)				
$PRI(t)SEC(t)^{-1}$			.003	−.005		
			(.002)	(.008)		
$TER(t)SEC(t)^{-1}$			.023	.412		
			(.075)	(.715)		
$PRI(t)TER(t)^{-1}$					−.003	.006
					(.012)	(.009)
$SEC(t)TER(t)^{-1}$					.010	−1.75
					(.041)	(.336)
GNP^{-1}	−.049	.084	.130	.128	1.59	2.41
	(.014)	(.128)	(.099)	(.163)	(4.44)	(8.84)
POP^{-1}	.090	.555	.525	−.608	2.03	−26.6
	(.056)	(.491)	(.208)	(1.70)	(9.61)	(82.7)
$PRIPOP^{-1}$	.140	−.701				
	(.081)	(.633)				
GNP^{-2}	.00002	−.00003				
	(.00002)	(.00001)				
Constant	.001	−.002	−.007	−.021	−.061	16.9
	(.002)	(.021)	(.015)	(.115)	(.396)	(4.56)
R^2	.973	.791	.917	.931	.912	.835
N	32	32	32	32	32	32
NT	128	128	128	128	128	128

NOTE: Standard errors are shown in parentheses.

Table 2. Estimates of Coefficients of Differential Equation Model (Equation 3) and Carrying Capacity Function (Equation 9)

	Dependent variable					
	$\frac{dPRI}{dt}$ (j=1)		$\frac{dSEC}{dt}$ (j=2)		$\frac{dTER}{dt}$ (j=3)	
	Rich	Poor	Rich	Poor	Rich	Poor
Independent Variables (Parameter):						
Speed of adjustment (a_j)	.061	.238	.169	.115	.101	.054
$PRI(b_{j1})$	†	†	−.001	.001	.001	−.001
$SEC(b_{j2})$	−.000	.005	†	†	−.003	.400
$TER(b_{j3})$	.000	−.029	−.007	−.108	†	†
Carrying capacity variables (Parameter):						
$GNP^{-1}(c_{j1})$	−.185	.122	.228	.292	4.01	10.2
$POP^{-1}(c_{j2})$	.340	.799	.919	−1.39	5.12	−113.
$PRIPOP^{-1}(c_{j3})$	.529	−1.009				
$GNP^{-2}(c_{j4})$	.00008	−.00004				
Constant(c_{j0})	.004	−.003	−.012	−.048	−.154	71.5

† Parameters $b_{jj} = -a_j(c_{j0} + c_{j1}GNP^{-1} + c_{j2}POP^{-1} + \theta c_{j3}PRIPOP^{-1} + \theta c_{j4}GNP^{-2})$ by definition where $\theta = 1$ when $j = 1$ and 0 otherwise.

of adjustment parameter a_j decrease monotonically from PRI to TER with one exception—the PRI coefficient (a_1) for rich is lower than expected. The general result agrees with the NH hypothesis (Proposition 3) concerning rates of equilibrium adjustment across organizational forms; the least costly and least complex form adjusts to changes in carrying capacity most rapidly. Also, again with the exception mentioned above, the adjustment speed parameters for each organizational form are higher in the rich nations. This supports the NH argument given in Proposition 4. The unexpected low value of a_1 for rich nations is probably due to an observation problem: the most significant expansion of primary systems in rich nations occurs prior to 1950 and is not captured by the data. Consequently, the growth rate for these organizations appears much slower.

The b_{jk} coefficients of the endogenous matrices are not so systematically patterned. Parameters of lower level educational systems are not always symbiotic as the NH analysis suggests. Parameters of the higher level systems are negative in four out of six instances, thereby suggesting competition. In general, though, these data do not allow adjudication between the two opposing effects. However, when the matrices of endogenous coefficients are compared across samples, the larger absolute values for the coefficients of the poor countries support the prediction of increased relational intensity in scarce environments (Proposition 2).

The estimates for the GNP^{-1} and POP^{-1} variables in Table 2 are also not systematically patterned. GNP^{-1} has a positive effect except when it is estimated with a second-order term. POP^{-1} has a positive effect in the PRI equations, but exhibits reversing signs across samples in the SEC and TER equations. Both variables usually have higher absolute values in the poor environment. However, interpretation of these coefficients is not straightforward, since they are nonlinearly embedded in the model. In order to understand these effects more intuitively, I have constructed Table 3.

In Table 3, I have used the definition of b_{jj} to compute the values of these parameters for given levels of the exogenous variables. I have also calculated the corresponding carrying capacity of each form when the other forms are not present. As expected, the carrying capacities decrease monotonically from PRI to TER for the rich subsample. For the poor subsample, the PRI and SEC coefficients are both in their expected positions relative to TER. However, the carrying capacity for SEC is slightly higher than the same parameter for PRI, and the feedback parameter for SEC is slightly smaller than expected.

The model behaves very reasonably when the 1950 exogenous variables deviate from the means. As expected, the carrying capacities are dramatically lower for nations with low levels of the exogenous variables. Moreover, the feedback parameters are much larger when the exogenous variables are at low levels. This observation suggests that growth is more constrained in the resource-scarce environments. This argument is also supported by the estimates computed using the 1965 (higher) mean levels of the exogenous variables. In all equations, these levels predict, as expected, smaller feedback parameters and increased carrying capacities.

Comparing these estimates across the rich and poor subsamples, we observe in the PRI equations, as expected, a higher carrying capacity for rich nations and a larger feedback parameter for poor nations. In SEC the reverse is true, although the differences are quite small. In the TER equations, the expected pattern is again obtained.

Figure 1 displays the estimated relationship between GNP and carrying capacity for primary enrollments when the model is estimated for poor nations. Since the constant plays an important role in the carrying capacity function (Equations 1), I have set the other exogenous variables at their 1950 means to calculate the points in this figure. As the graph shows, the estimated equation conforms to the prediction in Proposition 1: GNP increases the carrying capacity at decreasing rates. As might be expected given the discrepancy in the estimates reported above, this proposition is supported by the estimated equations for all but the PRI rich nation equation.

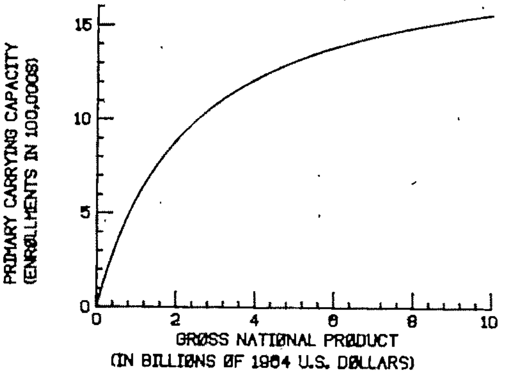


Figure 1. Estimated GNP Effect for the Poor Nations

STABILITY ANALYSIS

In Proposition 5 I argue that the expansion process in national systems of education will follow a logistic growth curve. This proposition highlights the main difference between my model and the NH model: the dynamic structure. However, on this point the NH model is weakest. In four out of the six equations, the favored estimator of the NH model yields "substantively meaningless dynamics" (NH: 489).

The plausibility of the dynamics of my model can be assessed by evaluating its "stability".[8] The stability of equations in a deterministic system is usually assessed at some particular realized point. *Stability* is the case where the system returns to the point after the variables have been perturbed initially. An *unstable* system is characterized by amplification away from the point when the exogenous shock hits. And finally, the system is *neutrally stable* if the oscillations caused by the shock continue with constant amplitude.

Stability is characterized according to the region to which it applies. The distinction between global and neighborhood (or local) stability is useful. Global stability assesses system behavior after perturbation from any possible point. Neighborhood stability evaluates the behavior only in the region near the equilibrium point. When the dynamic system is linear, global and local stability are exactly similar. However, for nonlinear models such as Equations 3, this is not necessarily the case since the time trajectory potentially contains many inflection points. In this case, analysis of global stability is quite complicated and often mathematically intractable (see May, 1973 and Siljak, 1978). Instead, only local stability can be analyzed.

Assessment of local stability is partially a methodological evaluation but it is also substantively informative. One of the most important theoretical claims of current organizational research is the prediction of isomorphism between organization and environment. Isomorphism, however, is an equilibrium prediction and is often not linked to the dynamic process by which it is obtained. It is important to know if nonequilibrium models such as the one proposed here reach an equilibrium state compatible with their internal logic and the prediction of isomorphism.[9] In our case, this includes the presence of all organizational forms and their viability relative to environments.

The model presented here is deterministic and assumes that all relevant exogenous variables are included in the specification or summarized by the disturbance. Because many idiosyncratic events such as energy shortages, political regime turnovers, and financial crises also affect the growth processes of national systems of education, it is important to know how resistant the equilibrium is to random environmental disturbances. If we think of these events as exogenous shocks to the differential equation system, then we can partially assess their effect upon the predicted trajectories.

I assess the stability of the model when the feedback parameters are constant at the values predicted by the 1950 mean

[8] I have chosen the stability analysis rather than the more traditional method of evaluating models by their fit to the data because the original NH model misspecifies the theoretical arguments, the estimates of the NH model yield an implausible long-term general model, the main point in contention concerns the time trajectory over long periods of time and the available panel data contain only a few points over a relatively short period, and the previous empirical evidence demonstrates that when long-term observations are available, the growth process appears logistic.

[9] Blalock (1969:77) argued the more general case that "an adequate theoretical explanation in causal terms requires an understanding of . . . stability conditions."

Table 3. Characteristic Values of Feedback Parameters (b_{jj}) and Carrying Capacity for Certain Levels of Exogenous Variables

	PRI			
	Rich		Poor	
	b_{11}	PRI*	b_{11}	PRI*
Values of exogenous variables:				
GNP POP PRIPOP				
Mean values in 1950	−.0103	5.922	−.1190	2.000
One standard deviation below 1950 mean	−.0227	.3721	−.2319	1.026
Mean values in 1965	−.0094	6.489	−.0642	3.707
	SEC			
	Rich		Poor	
	b_{22}	SEC*	b_{22}	SEC*
Values of exogenous variables:				
GNP POP				
Mean values in 1950	−.1492	1.133	−.0560	2.053
One standard deviation below 1950 mean	−.3712	.4553	−.1975	.5823
Mean values in 1965	−.0953	1.773	−.0115	10.00
	TER			
	Rich		Poor	
	b_{33}	TER*	b_{33}	TER*
Values of exogenous variables:				
GNP POP				
Mean values in 1950	−1.065	.0948	−3.509	.0154
One standard deviation below 1950 mean	−2.698	.0374	−3.899	.0138
Mean values in 1965	−.6472	.1561	−3.141	.0172

levels of the exogenous variables (see Table 3). Estimates of the equilibrium levels PRI**, SEC**, and TER** are then obtained by setting the growth equations (Equations 3) equal to zero and solving algebraically. I now expand around this equilibrium as follows:

$$\mathrm{PRI}(t) = \mathrm{PRI}^{**} + \chi_P(t) \quad (11a)$$
$$\mathrm{SEC}(t) = \mathrm{SEC}^{**} + \chi_S(t) \quad (11b)$$
$$\mathrm{TER}(t) = \mathrm{TER}^{**} + \chi_T(t) \quad (11c)$$

where χ_j indicates an initially small exogenous shock. If we expand the original Equations 3 by Taylor series around this equilibrium, we can discard the nonlinear terms (because if χ is close to zero, χ^2 is minuscule) and obtain the linear approximation

$$\frac{d\chi_j}{dt} = a'_{jP}\chi_P(t) + a'_{jS}\chi_S(t) + a'_{j}\chi_T(t) \quad (12)$$

of behavior around the equilibrium point. The set of parameters a'_{jk} has been defined as the "community matrix" by Levins (1968) and describes the interactional relations of form j on k near equilibrium. We can find these values by calculating the Jacobian of Equation 3, which I call the matrix A, as

$$a'_{jk} = (\partial F_j/\partial E_k)^* \quad (13)$$

where F_j are the functions given in Equations 3 and E_1 = PRI, E_2 = SEC and E_3 = TER. By evaluating these derivatives at the equilibrium levels we can determine the entries a'_{jk} for all j and k.

System behavior in this region can be evaluated by the Routh-Hurwitz criteria, which are the necessary and sufficient conditions of stability (Routh, 1905; Hurwitz, 1895). These criteria are applied by expanding the characteristic equation of the matrix

$$\det |A - \lambda I| = 0 \quad (14)$$

as a polynomial in λ. Specifically for the three-dimensional case this becomes

$$f(\lambda) = \lambda^3 + C'_1\lambda^2 + C'_2\lambda + C'_3 = 0 \quad (15)$$

where C'_1, C'_2 and C'_3 are constant functions of entries in A as given by (14). The Routh-Hurwitz stability conditions for this 3×3 case are

$$C'_1 > 0 \quad (16a)$$
$$C'_2 > 0 \quad (16b)$$
$$C'_1C'_2 > C'_3 \quad (16c)$$

Table 4 provides the estimates obtained in the stability analysis of the model with b_{jj} as determined by mean 1950 exogenous variables. For both samples, the equilibrium values are all positive and substantively meaningful. Moreover, in both cases the community matrices are well-behaved and the Routh-Hurwitz criteria show stability in the equilibrium regions. These findings suggest that exogenous shocks to the systems do not alter the previously reported results in the vicinity of equilibrium.

Table 4 presents the analysis of system behavior around the equilibrium point but it does not indicate whether the observed enrollment levels lie within the region moving towards this equilibrium. That is, the system could be locally stable, but the observed point could lie within an unstable region. To answer this question I have taken an extremely empirical approach and plotted the time trajectories predicted by the estimated equations. Figures 2 and 3 show the time paths for the rich and poor subsamples. In both instances, the 1950 mean observations move toward the equilibrium solutions. Also, as we might expect, these equilibrium levels are higher in the rich environment.

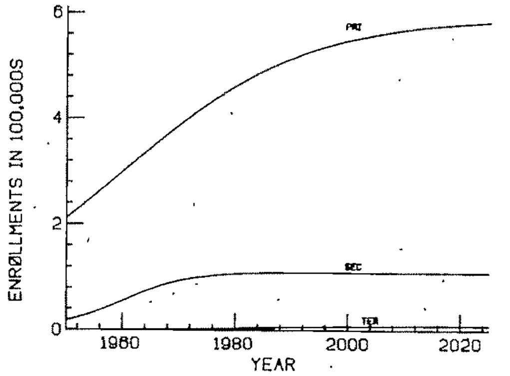


Figure 2. Predicted Time Path for Rich Nations.

This analysis supports the plausibility of the estimates obtained earlier. Nonetheless, the general expansion model estimated does not have constant b_{jj} parameters. The dependence of these parameters on the exogenous variables makes even local stability analysis intractable (and probably impossible). This difficulty arises in part due to the impossibility of finding the equilibrium solution when the feedback parameters change in an unspecified way.

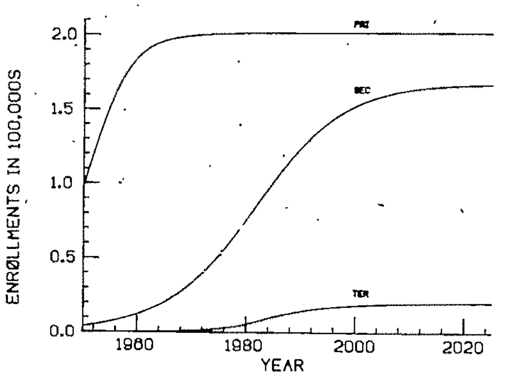


Figure 3. Predicted Time Path for Poor Nations.

Table 4. Estimates of the Community Matrix and the Coefficients of the Characteristic Polynomial

Dependent variable	Equilibrium value	Community matrix a'_{j1}	a'_{j2}	a'_{j2}	Coefficients of characteristic polynomial C'_1	C'_2	C'_3
			Rich Nations				
PRI(j=1)	5.922	−.061	0.0	0.0	.326	.032	.001
SEC(j=2)	1.088	−.001	−.162	−1.159			
TER(j=3)	.096	.001	−.0004	−.1025			
			Poor Nations				
PRI(j=1)	2.022	−.241	.008	.040	1.063	.282	.020
SEC(j=2)	1.688	.002	−.094	−.182			
TER(j=3)	.207	−.0002	.083	−.728			

DISCUSSION

I begin by reexamining NH's original theoretical claims. In Propositions 3 and 4, NH's arguments about organizational adjustment speeds to changed carrying capacities were reiterated. I claimed that the less complex and the less costly organizations would adjust more quickly. I also argued that, for similar organizational forms, the adjustment speeds would be higher in resource-abundant environments. In my judgment, the findings support both arguments and suggest the merit of future research along these lines.

NH's second major theoretical claim, which I did not reproduce here, concerned the degree to which organizational boundary-spanning units insulated the core technology from environmental uncertainty. With reasoning similar to Thompson (1967), they argued that the symbiotic cohort effects of lower level organizations would be greatest in the least complex and least costly organizations since there was less to protect. Rather than limit the analysis to this single type of interorganizational relation, I also allowed competition from other organizations to occur. Opening up the model in this way produced confusing results. NH's pattern of results disappeared and an alternative systematic pattern failed to emerge. I interpret this result as the consequence of simultaneous but opposing relationships between the organizational forms. When lower level educational systems expand, they enhance the conditions for expansion at the higher levels since they produce qualified students. But expansion of lower level systems also diminishes the chances for successful expansion at higher levels, since all levels make claims on the same finite set of resources and the prior expansion of one level shrinks the amount of resource available to others. With the present data, I am unable to discriminate between these processes although I believe more refined data might allow adjudication.

We might ask then, of what value is the model which allows any and all types of interorganizational relations? Its primary contribution here has been with respect to the relational intensity between organizational forms. The results show, as predicted by Proposition 2, that the intensity of links within the network of organizational forms is greatest in resource-scarce environments. Hannan and Freeman (1978) present similar results for declining versus growing educational organizations.

In my first proposition I argued that resources would have nonlinear effects on the carrying capacities of organizations. My reasoning was simple: when resources are at high levels, organizational carrying capacities are less sensitive to a unit increase than when resources are at low levels. The empirical findings displayed this relationship most clearly for GNP in the primary systems. It now appears that the proposition holds primarily for financial resources and for relatively low-cost, noncomplex organizations. Since these organizations adjust fastest to changes in carrying capacities, it also makes sense that they will be most sensitive to resource fluctuations.

The most dramatic departure between my model and NH's is my specification of a logistic expansion process. Somewhat ironically, I argued that the logistic more accurately specifies their own theoretical arguments than the linear partial adjustment used. Additional theoretical reasons for its use include the belief that new small organizations face many obstacles of legitimation, resource claims, and implementation that must be overcome before large-scale expansion can occur. Stinchcombe (1965), in a well-known discussion about the "liability of newness" in young firms, described similar problems with respect to internal coordination and consumer acceptance. Obstacles to organizational growth such as these may well be ubiquitous, but I think they are compounded when the organization is state-owned and geographically decentralized. This pattern of slow initial growth is well captured by the logistic specification, and the findings support my model over NH's. Whereas the estimates from their model lead to nonsensical long-term growth processes, the estimates from mine yield meaningful predictions even when subjected to a rigorous analysis of stability. These results provide considerable power

to the proposals of Hannan and Freeman (1977) and Aldrich (1979) to use the logistic-based Lotka-Volterra equations for the study of organizational growth and ecology.

Finally, the Lotka-Volterra model is considerably more general than I have yet acknowledged. The system of Equations 3 is a general dynamic model of interrelated variables used widely in bioecology, economics, and engineering (see Siljak, 1978); its main advantages are generality and flexibility. Current sociological applications are to the theories of organizations (Hannan and Freeman, 1977) and of ethnic groups (Hannan, 1979; Nielsen, 1978; 1980). Empirical work in these areas, however, has used only the linear partial adjustment model as an approximation. The techniques used in this study might be a more feasible estimation strategy.

REFERENCES

Aldrich, Howard E.
1979 Organizations and Environments. Englewood Cliffs, N.J.: Prentice-Hall.

Ayala, Francisco J., Michael E. Gilpin and Joan G. Ehrenfeld
1973 "Competition between species: theoretical models and experimental tests." Theoretical Population Biology 4:331–56.

Baloff, Nicholas
1971 "Extensions of the learning curve: some empirical results." Operations Research Quarterly 22:329–40.

Blalock, Hubert M.
1969 Theory Construction. Englewood Cliffs, N.J.: Prentice-Hall.

Brittain, Jack and John Freeman
1980 "Organizational proliferation and density-dependent selection." Pp. 291–338 in J. Kimberely and H. Miles (eds.), Organizational Life-Cycles. San Francisco: Jossey-Bass.

Etzioni, Amitai
1968 The Active Society. New York: Free Press.

Glassow, Michael A.
1978 "The concept of carrying capacity in the study of culture process." Advances in Archaelogical Method and Theory 1:31–48.

Hamblin, Robert, Ralph B. Jacobsen, and Jerry L. Miller
1973 A Mathematical Theory of Social Change. New York: Wiley.

Hannan, Michael T.
1979 "Dynamics of ethnic boundaries." Pp. 253–75 in John Meyer and Michael T. Hannan (eds.), National Development and the World System: Educational, Economic and Political Change, 1950–1970. Chicago: University of Chicago Press.

Hannan, Michael T. and John Freeman
1977 "The population ecology of organizations." American Journal of Sociology 82:929–66.
1978 "Internal politics of growth and decline." Pp. 177–99 in M. Meyer and Associates (eds.), Environments and Organizations. San Francisco: Jossey-Bass.

Hannan, Michael T. and Nancy Brandon Tuma
Forthcoming Social Dynamics: Models and Methods. New York: Academic Press.

Hannan, Michael T. and Alice Young
1977 "Estimation in panel models: results on pooling cross-sections and time-series." Pp. 52–83 in David Heise (ed.), Sociological Methodology 1977. San Francisco: Jossey-Bass.

Hawley, Amos
1950 Human Ecology. New York: Ronald Press.

Henderson, Charles R.
1952 "Specific and general combining ability." Pp. 352–70 in J. W. Gowen (ed.), Heterosis. Ames, Iowa: Iowa State College Press.
1963 "Selection index and expected genetic advance." Pp. 141–62 in Statistical Genetics and Plant Breeding. Publication No. 982. Washington, D.C.: National Academy of Sciences.

Hurwitz, A.
1895 "Uber die Bedingungen unter welchen eine Gleichung nur Wurzeln mit negativen reellen Teilen besitzt." Mathematische Annalen 46:273–84.

Hutchinson, Evelyn
1979 An Introduction to Population Ecology. New Haven: Yale University Press.

IBRD (International Bank for Reconstruction and Development)
1971 World Tables. Washington, D.C.: IBRD.

Leslie, P. G.
1958 "A stochastic model for studying the properties of certain biological systems by numerical methods." Biometrika 45:16–31.

Levins, Richard
1968 Evolution in Changing Environments. Princeton: Princeton University Press.

Lipset, Seymour Martin
1960 Political Man. Garden City, N.Y.: Doubleday.

May, Robert M.
1973 Stability and Complexity in Model Ecosystems. Princeton: Princeton University Press.

Meyer, John W.
1977 "The effects of education as an institution." American Journal of Sociology 83:55–77.

Meyer, John W., Michael T. Hannan
1979 National Development and the World System: Educational, Economic and Political Change, 1950–1970. Chicago: University of Chicago Press.

Meyer, John W., Francisco Ramirez, Richard Rubinson, and John Boli-Bennett
1977 "The world educational revolution." Sociology of Education 50:242–58.

Nielsen, François
1978 "Linguistic conflict in Belgium: an ecologi-

cal approach." Ph.D. dissertation. Stanford University.
1980 "The Flemish movement in Belgium after World War II: a dynamic analysis." American Sociological Review 45:76–94.
Nielsen, François and Michael T. Hannan
1977 "The expansion of national education systems: tests of a population ecology model." American Sociological Review 42:479–90.
Pemberton, H. Earl
1936 "The curve of cultural diffusion." American Sociological Review 1:547–56.
Pianka, Eric C.
1978 Evolutionary Ecology. New York: Harper and Row.
Pielou, E. C.
1977 Mathematical Ecology. New York: Wiley and Sons.
Pressman, Jeffrey and Aaron Wildavsky
1973 Implementation. Berkeley: University of California Press.
Rogers, Everett with F. Floyd Shoemaker
1971 Communication of Innovations. New York: Free Press.
Rosenfeld, Rachel
1980 "Race and sex differences in career dynamics." American Sociological Review 45:583–609.
Routh, Edward J.
1905 The Advanced Part of a Treatise on the Dynamics of a System of Rigid Bodies. London: Macmillan.
Siljak, Dragoslav D.
1978 Large-scale Dynamic Systems: Stability and Structure. New York: North Holland.
Stinchcombe, Arthur
1965 "Social structure and organizations." Pp. 153–93 in James G. March (ed.), The Handbook of Organizations. Chicago: Rand McNally.
Taylor, Charles L. and Michael C. Hudson
1971 World Handbook of Political and Social Indicators, II. Ann Arbor: Inter-University Consortium for Political Research.
Thompson, James D.
1967 Organizations in Action. New York: McGraw-Hill.
Thompson, James D. and William McEwen
1958 "Organizational goals and environment: goal-setting as an interaction process." American Sociological Review 23:23–31.
UNESCO (United Nations Educational, Scientific and Cultural Organization)
1971 Statistical Yearbook. Louvain, Belgium: UNESCO.

COMMUNITY STRUCTURE IN RESPONSE TO POPULATION GROWTH AND DECLINE: A STUDY IN ECOLOGICAL ORGANIZATION*

GLENN V. FUGUITT
University of Wisconsin–Madison

JOHN D. KASARDA
University of North Carolina–Chapel Hill

Building upon community and organizational theory, we assess how three key components of community structure—managerial, clerical, and professional and technical support—respond to population size, growth, and decline. The units are 1,395 relatively autonomous U.S. nonmetropolitan cities. For places that grew between 1960 and 1970, our hypotheses were confirmed that a positive association exists between both population size and growth and the proportion of employed in professional and clerical occupations. A negative association for the managerial component is consistent with findings in organizational research. Smaller but similar associations were generally found for declining places. Status of communities as administrative centers and population change in their hinterlands are introduced as additional variables. Alternative longitudinal models yield consistent results: structural changes tend to be symmetrical across growing and declining communities, but they are less responsive to demographic contraction than expansion.

Modern human ecologists define community structure as the interdependent complex of sustenance functions through which the resident population maintains itself (Duncan and Schnore, 1959; Frisbie and Poston, 1975, 1978; Gibbs and Martin, 1959; Hawley, 1950, 1968). Structural analysis of the community primarily involves three tasks: identifying the significant functions of the sustenance complex; determining the morphological pattern of the complex, i.e., the relative numbers and kinds of functions and their relationships (Hawley, 1950:206–33); and assessing the consequences for the sustenance organization of change in pertinent demographic, environmental, and technological factors.

Though most ecological theories of community structure and change are dynamic, the bulk of empirical analyses are static (cf. Duncan and Reiss, 1956; Frisbie and Poston, 1975, 1978; Kasarda, 1972; Smith and Weller, 1977). Although cross-sectional analyses that meet the necessary assumptions can inform underlying structural dynamics, confidence in findings increases when theory and measurement harmonize (Blalock, 1979; Hannan and Tuma, 1979; Frisbie, 1980).

Our research employs longitudinal analysis of 1,395 nonmetropolitan communities to examine the consequences of population size, growth, and decline for the sustenance structure of communities. (See Spencer, 1877; Durkheim, 1893; and Halbwachs, 1938, for early discussions of this issue.) Analysis of the static relationship between the size and form of communities will be augmented by dynamic analysis

* Direct all correspondence to Glenn V. Fuguitt, Department of Rural Sociology, University of Wisconsin, 240 Agricultural Hall, 1450 Linden Drive, Madison, WI 53706.

Our names are listed in alphabetical order. This research has been supported by the College of Agricultural and Life Sciences, University of Wisconsin–Madison, and the Economic Development Division, U.S. Department of Agriculture, through a cooperative agreement; and by the Center for Demography and Ecology, University of Wisconsin–Madison, through a grant from the Center for Population Research of the National Institute of Child Health and Human Development. We are indebted to Dan Lichter, Tim Heaton, and Mary Jacobs for their assistance throughout all phases of the analysis, and acknowledge also helpful comments from David Heise, Keith Warner, and anonymous referees of the *American Sociological Review*.

of the relationship between population growth and structural change. The communities are divided between those that gained and lost population over a ten-year period, to permit us to assess whether changes in sustenance organization for growing and declining communities are equivalent.

Our analytical framework is derived from Kenneth Boulding (1953). Synthesizing the pioneering work of biologist D'Arcy Thompson (1917) on the relationship between the growth and form of organisms with later work by economists on compensatory change, Boulding describes basic growth processes in social systems and their consequences for internal structure.

Boulding begins by explicating the distinction between *simple* growth and *structural* growth in social systems. Simple growth involves increments or decrements in discrete quantities, elements, or parts of the system (e.g., residents, auto mechanics, government employees, tax revenue). Structural growth or decline involves changes in the proportions or other relations of the quantities, elements, or system parts to one another (e.g., government employees per resident).

Although conceptually distinct, simple growth and structural growth are functionally interdependent in the course of system development. Based on this dependency, Boulding provides general principles of structural development; the most germane for us is his principle of nonproportional change (Boulding 1953:335–6). Boulding argues that "As any structure grows, the proportions of its significant variables *cannot* remain constant." He draws an analogy between different sizes of organizational structure and different-sized models of a physical object. Uniformly increasing an object's linear dimensions squares its areas and cubes its volumes, with corresponding differences in structural variables dependent on linear, areal, or volumetric dimensions: a two-fold increase in linear dimensions increases volume eight times. Similarly, the elements of an institutional structure will grow in differing proportions. In particular, according to Boulding, increasingly large organizations experience increasing difficulty in maintaining communications between the executive center and the peripheral "surfaces" (e.g., classroom, parish, retail outlet). The surface grows proportionally more than the "linear" communication network can support, until compensatory increases must be made in the relative size of communications systems and other specialized administrative support functions.

RESEARCH HYPOTHESES

If Boulding's principle of nonproportional change is a general principle of structural adaptation to system size and growth, communities should develop proportionately larger administrative support structures as they grow in size. Moreover, if communication (information processing, duplication, and transmission) is *the* critical administrative problem that larger systems face as they expand, the largest impact of system size, growth, and decline should be on this component. Thus, larger communities should exhibit relatively larger proportions of workers who maintain and facilitate communications, and, over time, faster growing communities should witness disproportionately greater increments of workers performing these functions. Communities with larger declines in population size should reduce communication functions at a faster rate or increase them at a slower rate (relative to other system functions) than communities with smaller population losses.

Whereas communication may be central to holding together larger and growing social systems, the administrative structure of most social systems contains two other very important components: managerial and supervisory functions and professional and technical support functions (Blau, et al., 1966; Rushing, 1966; Haire, 1959; Stinchcombe, 1959). These two components may differ in their associations with system size, growth, and decline (Freeman and Hannan, 1975; Kasarda, 1974; Starbuck, 1965; Rushing, 1966; Thompson, 1969). The literature on organizations suggests that managerial and supervisory functions increase at a slower rate than the other sustenance functions in faster growing communities

because system size and growth result in increased functional segmentation, (i.e., incrementation of functionally equivalent roles) enabling increased spans of control (Blau, 1970) and hence lower managerial and supervisory ratios.

Community research, however, has yielded varied results. Winsborough (1960) found, for a set of large cities, that the proportion of managers and proprietors was negatively associated with city size, net of industrial composition. Although Kasarda (1974) found a similar negative association for school districts, he reported a very small positive association for Wisconsin small towns and a moderately positive association for nations. Erbe (1977), differentiating managers from proprietors, found the proportion of managers positively and the proportion of proprietors negatively associated with small city size.

If structural changes in managerial and supervisory functions are equivalent across growing and declining communities, the direction of association for declining places should be the same as that for growing places. With a negative association, for example, communities losing the most population should also experience the largest proportional increases (or smallest proportional declines) in their managerial and supervisory functions. This would be true even if there were an absolute decline in the number of managers and supervisors, as long as the number employed in all other activities declined faster.

A positive relationship would be expected between population size and the proportion in the professional and technical support structure of communities, as well as between population change and change in the professional and technical support structure. Large system size generates greater problems of information gathering, evaluation, and planning, which can be handled more efficiently by specialized professional and technical staff. Winsborough (1960), Kasarda (1974), and Erbe (1977) found support for this association with various units of analysis considered in each case at a single time.

If this hypothesis is to survive longitudinal analysis, we should observe positive associations between community population changes and changes in the proportion of workers performing professional and technical support functions. If the relationship is parallel for growing and declining situations, we could expect this positive association to hold for both growing and declining communities.

Considering the effects of population size/growth on the community's communication, managerial, and professional and technical support structure, we hypothesize that all relationships between the total (combined) administrative and support structure and population size, growth, and decline would be positive. This hypothesis, which follows from Boulding's (1953) principle of compensatory change, is based on the premise that any administrative economies of scale in the managerial and supervisory sector would be more than compensated for by diseconomies (i.e., greater than proportional increases) in the communication, professional, and technical support sectors as community systems grow, and vice versa as communities decline.

We also hypothesize that the overall administrative structure will respond similarly in both growing and declining communities. If such is the case, communities that experience the largest declines in population should also experience the largest declines in their overall administrative proportion. Several studies have indicated, however, that changes in the administrative ratio should be different in declining situations as contrasted with growth (Tsouderos, 1955; Haire, 1959; Hendershott and James, 1972; Freeman and Hannan, 1975). The most intriguing question from this research is whether administrative "featherbedding"—the retention of "unnecessary" managerial, professional, or technical personnel—appears in a declining situation. A more charitable interpretation would be that once formal procedures are established for improving coordination and communication, they may not be easily reduced or dismantled.

We have chosen an additional attribute that might affect a community's structural response to population change: whether

the place is a county seat. (See Christaller's (1966) distinction between administrative and simple market centers.) The expansion of local government activities during the 1960s was felt in rural county-seat towns. Whether these expanding administrative responsibilities result in a different relationship between population change and change in the proportion employed in administrative occupations is not easy to predict. In growing situations perhaps the expansion of administrative activity in government and related private activity, such as law practice, would lead to a larger association between administrative proportion and population size and change than for non–county seats. In declining situations, administrative "featherbedding" or even increased welfare needs, if decline reflects a depressed economic situation, might lead to an increased administrative proportion.

The special consideration of county-seat status implies an open rather than a closed community system perspective within the county. We have tried to restrict our attention to relatively autonomous communities, but these places, and particularly the county seats, clearly should not be experiencing changes independent of changes in the county as a whole. In the growing case, particularly in metropolitan areas, this has been expressed in terms of the theory of ecological expansion (McKenzie, 1933; Hawley, 1950; Kasarda, 1972) viewed as a progressive absorption of more or less unrelated populations into a single organization. One central proposition of this theory holds that an increase in relative administrative activity in the center is a consequence of population expansion in the periphery. We examine this proposition for places, classified separately by county-seat status and place growth and decline, comparing the association of change in the administrative structure with change in (a) the population of this place, (b) total county population, including this place, and (c) county population outside of this place.

DATA MEASUREMENT AND METHODS

The community units for this study are all of the 1,395 U.S. nonmetropolitan incorporated places having more than 3,000 people in 1960 and more than 2,500 people in 1970, in relatively autonomous counties. These are nonmetropolitan counties more than 50 miles from metropolitan centers as of 1963, excluding rural counties from which more than 10% of the population commuted to metropolitan or more urban nonmetropolitan counties in 1960. Because data were not available for places with less than 2,500 population at either census date, the universe was set initially at slightly greater than 2,500 to allow observation of places declining in population. Of those over 3,000 in 1960, only three had fewer than 2,500 in 1970 and so had to be excluded. The county classification was a combination of the SMSA (Standard Metropolitan Statistical Area) adjacency code developed by Hathaway, Beegle, and Bryant (1968) and a location code by Calvin Beale. The latter classifies nonmetropolitan counties as (1) "trade center" counties: those having either an urban population of 25,000 or more, or 10,000 or more nonagricultural jobs; (2) commuter counties, from which more than 10% of the employed labor force commuted to "trade center" or SMSA counties in 1960; and (3) other counties, which would be more rural but relatively autonomous, at least in terms of labor-force participation. We excluded incorporated places located in counties within 50 miles of the central city of an SMSA and those others in counties classed by Beale as commuter counties.

We selected places with low interdependence with metropolitan areas and low rates of commutation for three reasons. First, because of their greater relative autonomy (before 1970 at least), our universe of places approximates the traditional ecological representation of "community." Second, because of their autonomy from metropolitan labor markets, we can analyze the structural responses of our communities to their own demographic changes better than those of communities more dependent upon metropolitan centers for local managerial communication and professional and technical support functions. Third, *de jure* census data allocate occupational activities by place of residence rather than by

where the activity is performed. Hence, the census occupational structure of communities situated in areas of high commutation would not reflect adequately the internal activities of the communities.

Of course, as Hawley points out (1950:222–3), true independence exists only for completely isolated and self-sufficient communities. Though this is not the case in any segment of nonmetropolitan America, absolute autonomy is not necessary for the theory (though its lack may make alternative explanations more salient), nor does absolute autonomy characterize the units used in previous parallel research on organizations or societies.

Our index of the overall administrative structure of communities as drawn from the 1960 and 1970 Census of Population is the proportion of all employed residents performing administrative functions[1]—residents in the census categories of proprietors, managers, and officials; professional, technical, and kindred; and clerical and kindred workers. Though the census categories represent far less than ideal measurement of "administrative statuses," they are the most appropriate data available and are either identical or comparable to most measurements of administrative statuses in the related formal organization research (see Stinchcombe, 1959; Rushing, 1966). (Note that sales workers, the remaining Census white-collar category, is not included as administrative.)

The census data have the further advantage of allowing us to decompose the total administrative structure roughly into its managerial and supervisory, communications, and professional and technical auxiliary staff components. Our index of managerial and supervisory structure is the proportion of all economically active residents in the community classified as managers, officials, and proprietors except farm (1960) or managers and administrators except farm (1970), to single out those involved in such activities as decision-making, supervision, coordination, and planning.[2]

Our index of communication structure is the proportion of workers employed in clerical and kindred activities. Clerical personnel are the ones most directly concerned with communication (recording, duplicating, and transmitting information) in modern bureacratic systems (Blau et al., 1966; Haire, 1959). These individuals serve to maintain communication through keeping written and machine records, channeling telecommunications, and arranging face-to-face meetings.

Our index of professional and technical support structure is the proportion of workers classified as professional, technical, and kindred workers. Professional and technical staff make contributions to overall administration by detecting problems and/or proposing knowledgeable solutions for them. However, the effectiveness of the professional and technical staff depends, in turn, on its being complemented by adequate managerial and clerical components. "These are required to meet the problems of coordination and communication that professionalization raises, notwithstanding the ability of professionals to help solve these problems" (Blau et al., 1966:191). Thus we see that the three basic components of the overall administrative fabric of both formal organizations and communities, while analytically distinguishable, are interdependent.

For cross-sectional analysis of the relationship between the size and form of communities, we use two baseline models. The first is $Sy = a + b\,(\ln P)$, where Sy refers to the structural property y of the community (proportion in administrative component y) and P the population size (logarithmically transformed). The formal rationale for this model is provided by Mayhew and Levinger (1976). Our second

[1] This index follows the standard sociological definition of structural properties of social systems as "properties of collectives which are obtained by some operation on the data about the relations of each member to some or all of the members" (Lazarsfeld and Menzel, 1961).

[2] For the managerial category, as well as for the professional and technical and clerical and kindred categories, there are minor differences between 1960 and 1970 due to classification changes for some individual occupations. The effect must be small, however, as a study showed discrepancies of 4.0%, 3.4%, and 1.9% when the U.S. 1960 populations of these three categories are classed according to 1970 procedures (Priebe, Heinkel, and Greene, 1972:2).

cross-sectional model takes the form of Y = aP^b, where Y is the absolute number in the administrative component y. If one logarithmically transforms this equation, it becomes log Y = log a + b log P, which is an allometric model indexing the relationship between the size of a system and any one of its parts (Svalastoga, 1974). The slope (b coefficient) in this model is the same as that which economists refer to as the elasticity coefficient (see Johnston, 1972). It indicates the percentage change in a given structural component corresponding to a one percent change in the size of the system.

Dynamic analysis presents a number of difficulties; the most troublesome of which is autocorrelation in the presence of measurement error. (See Bohrnstedt, 1969; Cronbach and Furby, 1970; Bereiter, 1962; Hibbs, 1974; and Ostram, 1978. For an overview of longitudinal measurement see Hannan and Tuma, 1979.) We present the results from five dynamic models specified in different manners, ranging from simple first differences to residualized change models containing both lagged endogenous and lagged exogenous variables.

RESULTS

Cross-Sectional Analysis

Our assessment of the relation between population size/change and change in community structure begins with a separate examination of results for the two census dates 1960 and 1970, considering separately those communities that grew or declined over the interval.

For baseline descriptors, the 1960 and 1970 mean values of the variables for the groups of nonmetropolitan places to be considered are provided in Table 1. The first two rows show that the number employed in all administrative occupations is slightly more than one third of the total employed labor force, about equally distributed among the professional, managerial, and clerical subcategories. The total administrative proportion was larger in 1970 than in 1960 for all classifications of communities, with increases in professional and clerical proportions more than offsetting the decline in the managerial proportion.

Comparison of the types of communities shows that growing places have somewhat higher administrative proportions than those that have declined, con-

Table 1. Mean Proportions of the Employed Labor Force in Administrative Categories for Types of Nonmetropolitan Urban Places

Types of places and year	(N)	Administrative	Professional	Managerial	Clerical
All places					
1960	(1395)	35.1	12.0	11.0	12.2
1970		37.0	13.5	9.4	14.1
Growing places					
1960	(868)	35.7	12.4	11.0	12.3
1970		37.8	14.1	9.5	14.2
County seats					
1960	(573)	36.3	12.5	11.4	12.4
1970		38.3	14.1	9.8	14.4
Not co. seats					
1960	(295)	34.6	12.2	10.4	12.0
1970		36.9	14.1	9.0	13.9
Declining places					
1960	(527)	34.1	11.2	10.8	12.1
1970		35.7	12.5	9.3	13.8
County seats					
1960	(325)	35.1	11.5	11.4	12.3
1970		36.6	12.9	9.8	13.9
Not co. seats					
1960	(202)	32.5	10.7	9.9	11.9
1970		34.2	12.0	8.6	13.6

Table 2. Correlations of Proportions of the Employed Labor Force in Administrative Categories with Log Population Size for Types of Nonmetropolitan Urban Places

Types of places and year	Administrative	Professional	Managerial	Clerical
All places				
1960	.15*	.10*	−.16*	.34*
1970	.32*	.27*	−.09*	.40*
Growing places				
1960	.21*	.16*	−.13*	.38*
1970	.39*	.33*	−.06	.45*
County seats				
1960	.21*	.17*	−.22*	.44*
1970	.41*	.36*	−.14*	.48*
Not co. seats				
1960	.17*	.14*	−.08	.25*
1970	.37*	.31*	.00	.36*
Declining places				
1960	.02	−.03	−.21*	.28*
1970	.11*	.06	−.15*	.29*
County seats				
1960	−.08	−.10	−.35*	.26*
1970	.07	.02	−.26*	.36*
Not co. seats				
1960	.05	.01	−.17*	.29*
1970	.08	.06	−.11	.19*

* Significant at the .05 level.

firming research of Ogburn (1937), Duncan and Reiss (1956), and Fuguitt and Field (1972). Similarly, county seats have higher proportions of their employed population in all three administrative categories, whether they are growing or declining.

There are positive associations between the log of population size and the total administrative, the professional, and the clerical proportions for each census date (see Table 2). There is a negative association between size and the managerial proportion. Correlations are quite low and nonsignificant, however, for the professional proportion and for the overall administrative proportion in declining places. County-seat status appears to have little influence on these zero-order correlations.

The fact that smaller places have higher proportions in managerial functions is due to the self-employed portion of this component, at least in 1970. For that year, using U.S. Census fourth count summary tapes, we were able to separate salaried from self-employed managers and administrators, and found that the −.09 for all places in Table 2 was a composite of .11 for salaried and −.33 for the self-employed. The proportion of all employed who are self-employed in retail trade correlated −.38 with log population size, and other self-employed −.14. The negative association between the managerial/proprietor proportion and log size thus appears to be due primarily to the greater importance in smaller, nonmetropolitan places of the retail trade function, particularly that represented by modest establishments.[3]

Comparing the two time periods, with one exception the 1970 correlations in Table 2 are more positive or less negative than those for 1960. That is, changes over the decade led to an increase in the size-proportion relationship for professionals and clericals, and the trend for the managerial proportion is in the same direction, with a smaller negative relationship.

As the second approach to cross-sectional analysis, the allometric model is useful for two reasons. First, the metric slopes (b coefficients) provide a straightforward indication of the percentage change in each administrative function with a one-percent change in population.

[3] This reflects the well-known fact that larger places are more likely to have other activities such as manufacturing, medical, and educational establishments in addition to their retail trade functions.

size. Second, a current methodological issue in sociology pertains to possible problems in correlating variables having common terms. This is not a problem here, since the administrative component is an intrinsic concept—a concept defined by its measure (see Schuessler, 1974; Fuguitt and Lieberson, 1974; Kasarda and Nolan, 1979; MacMillan and Daft, 1979). Nevertheless, since the denominator of the proportion of the employed labor force in an administrative category would be expected to correlate highly with log of population (correlations are generally above .90), the allometric model represents a parallel approach that avoids the common-term form.

Table 3 provides the results of the allometric analysis for the same categories and types of communities shown in Table 2. Here an exponent greater than one shows that larger places have a greater than proportional size of the component, and smaller places have a less than proportional size of the component, paralleling for the first model a positive correlation between total population and the proportion in the administrative category.

Tables 2 and 3 show closely parallel results regardless of method. Positive associations in Table 2 almost always correspond to coefficients greater than one in Table 3, and values are generally larger in 1970 than 1960. Similarly, negative associations found for managerial proportions in Table 2 usually correspond to coefficients less than one in Table 3, though these allometric coefficients are larger, and a few are equal to or greater than one by 1970.

These cross-sectional results for both methods for 1960 and 1970 are generally consistent with our expectations. The increase in size of these associations over the ten-year period indicates a strengthening of the hypothesized positive relationship for professional and clerical workers, but a weakening of the negative relationship for managers. The results do not confirm, however, the proposition that population change is associated with change in the administrative proportions over time.

Patterns of Change over the Decade

We first consider the absolute growth and decline of the population and the occupational categories between 1960 and 1970 (Table 4). The average population is

Table 3. Metric Slopes (b) for the Regression of Log Number in an Administrative Category on Log Population Size for Types of Nonmetropolitan Urban Places

Type of places and year	Administrative	Professional	Managerial	Clerical
All places				
1960	1.05*	1.06*	.97*	1.13*
1970	1.10*	1.13*	1.00	1.14
Growing places				
1960	1.07*	1.08*	.98	1.14*
1970	1.12*	1.17*	1.01	1.16*
County seats				
1960	1.05*	1.06*	.94*	1.13*
1970	1.10*	1.14*	.97	1.15*
Not co. seats				
1960	1.08*	1.10*	1.00	1.14*
1970	1.14*	1.20*	1.05	1.15*
Declining places				
1960	1.03	1.02	.96*	1.11*
1970	1.05*	1.06*	.97	1.11*
County seats				
1960	.99	.98	.90*	1.08*
1970	1.02*	1.02*	.92	1.11*
Not co. seats				
1960	1.06*	1.05*	.97	1.16*
1970	1.06*	1.08*	.97	1.11*

* Significantly different from 1.0 at the .05 level.

Table 4. Mean Population Size and Mean Percent Change in Absolute Size of Administrative Categories for Types of Nonmetropolitan Urban Places

Types of places and year	(N)	Administrative	Professional	Managerial	Clerical	Employed labor force	Total population
All places							
1960	(1395)	1,344	460	395	489	3,732	10,338
1970		1,681	631	398	652	4,295	11,452
Mean % change		22	33	2	36	15	11
Growing places							
1960	(868)	1,389	489	402	498	3,750	10,378
1970		1,926	743	447	737	4,762	12,664
Mean % change		35	47	11	50	26	22
County seats							
1960	(573)	1,564	547	456	562	4,169	11,408
1970		2,144	817	500	826	5,247	13,779
Mean % change		31	41	8	47	23	19
Not co. seats							
1960	(295)	1,050	377	298	374	2,937	8,379
1970		1,503	598	341	563	3,820	10,499
Mean % change		42	57	19	56	32	27
Declining places							
1960	(527)	1,270	413	383	474	3,702	10,270
1970		1,277	447	317	513	3,526	9,454
Mean % change		1	10	−14	13	−4	−8
County seats							
1960	(325)	1,444	468	443	533	4,130	11,385
1970		1,451	508	366	578	3,925	10,476
Mean % change		1	10	−16	12	−4	−8
Not co. seats							
1960	(202)	989	323	287	379	3,014	8,475
1970		999	349	241	409	2,886	7,810
Mean % change		2	10	−11	14	−4	−8

NOTE: These are the means of values obtained for each urban place.

slightly over 10,000 and the average number in the employed labor force is approximately 4,000. Since the distribution is skewed to the right of the mean, one would expect the median size of place and size of employed labor force to be somewhat smaller than these mean values.

Initial size differences between growing and declining places are negligible, though county seats tend to be somewhat larger than other centers. Moreover, there is no growth advantage evident for county seats during the decade. Among declining places, both county seats and non–county seats lost an average of 8%. On the other hand, among growing places non–county seats grew more rapidly than county seats. Though findings on the growth advantage for county seats are mixed (compare Fuguitt, 1965, with Tarver and Beale, 1968), here we see that whereas county seats are more likely to have grown over the decade, among growing places non–county seats grew relatively faster.

Overall, the number in the professional and clerical categories grew at three times the rate of the population and more than twice as rapidly as the number in the employed labor force, whereas the mean number of managers and administrators did not change over the decade. Growing places posted a nearly 50% growth in professionals and clericals, more than twice the rate of population growth, and managers grew at one-half the rate of population. Among declining places there also was a small absolute increase in professional and clerical workers, but the number of managers decreased absolutely at a rate almost twice that of the population decline. In Table 1, declining places showed an increased proportion of professional and clerical workers and a decrease for managers and administrators. The manager category may be particularly sensitive to population decline in part because of the decreasing importance of retail trade activity in smaller towns and cities. On the other hand, national trends toward increased professional, service,

and other system requirements lead to a slight increase in professional and public service employment despite population decline.

We begin our examination of the statistical associations between population change and changes in the labor force by considering the results from the following five regression models:

1. $AdP7 - AdP6 = b_1\,(LPop7 - LPop6) + a + e$
2. $AdP7 = b_1\,(LPop7 - LPop6) + b_2\,AdP6 + a + e$
3. $AdP7 = b_1\,LPop7 + b_2\,LPop6 + b_3\,AdP6 + a + e$
4. $AdP7 - \widehat{AdP7} = b_1\,(LPop7 - \widehat{LPop7}) + a + e$
5. $AdP7 = b_1\,(LPop7 - \widehat{LPop7}) + b_2\,AdP6 + a + e$

where AdP7 stands for an administrative proportion for 1970, AdP6 a proportion for 1960, LPop7 the population of the city in 1970, and LPop6 the population in 1960. The expression $(LPop7 - \widehat{LPop7})$ in Models 4 and 5 is the deviation of the observed 1970 log population of a place from that predicted on the basis of its 1960 population by a bivariate linear regression. One or another of these models has been endorsed by various writers, and several have been recommended as superior to the simplest, Model 1, the regression of first differences (Bohrnstedt, 1969; Bereiter, 1962; Heise, 1970). The five are actually quite similar in mathematical form, though leading to somewhat different least-squares estimations. We believe that if one posits continuous rather than a single discrete change over the decade, none of these models is necessarily inferior to others in terms of simultaneity bias. Because data are available at only two points in time, however, all are necessarily deficient in revealing the structure and timing of change. Rather than making a commitment here to any of these models as the best or most appropriate, we decided to present initial results from all five models of change.

The corresponding metric b-coefficients for the five models, showing the relationship between change in population and change in administrative proportion, (b_1 in each of the models above) are displayed in Table 5. The results are quite uniform across the models, and substantive conclusions are identical regardless of the model employed. As we hypothesized, growth in population is positively associated with growth in proportion employed in overall administrative, professional, and clerical occupations, and it is also negatively associated with growth in the managerial proportion. Coefficients for declining places are smaller and not significant in any case.

Because these results are quite similar, we report here only Model 4 in our more elaborated analysis. We selected Model 4 for two reasons. First, unlike Model 1, Model 4 has independent and dependent variables which correlate at zero with their respective initial levels (though we see that this makes little difference). Second, in contrast to Models 2, 3, and 5, our Model 4, as a simple bivariate regression, provides results in a form comparable to the cross-sectional analysis of Table 2. (In the bivariate form, Betas (standardized slopes) and correlation coefficients are equivalent.)

Table 6 provides analogous results to Table 5, but places are differentiated by whether or not they are county seats. For growing places the results are similar for both types of places, with all but one coefficient statistically significant and in the expected direction. The clerical coefficient here is larger and the professional coefficient smaller for county seats than non–county seats, but separate t-tests show neither difference is statistically significant.

For declining places, no coefficients are significantly different from zero, as was also true in Table 5. The coefficients of change for the clerical component here are considerably larger and the coefficients of change for the professional component are considerably smaller for non–county seats than for county seats. In fact, the professional component is negative for non–county seats, as is the administrative component.

Table 5. Corresponding Metric Slopes (b) and Standardized Slopes (Beta) for Six Models of Change in the Administrative Proportion Associated with Log Population Change, Nonmetropolitan Urban Places

	Administrative		Professional		Managerial		Clerical	
	b	Beta	b	Beta	b	Beta	b	Beta
All communities								
Model 1	2.24*	.11	2.38*	.18	−1.39*	−.11	1.24*	.10
2	3.37*	.10	2.81*	.12	−.91*	−.07	1.85*	.12
3	3.51*	.37	2.86*	.47	−.91*	−.27	1.93*	.46
4	3.04*	.15	2.52*	.19	−.91*	−.09	1.78*	.15
5	3.30*	.09	2.78*	.12	−.91*	−.07	1.82*	.12
Growing communities								
Model 1	2.50*	.11	3.28*	.23	−2.35*	−.18	1.57*	.12
2	3.67*	.09	3.63*	.13	−1.77	−.14	2.43*	.14
3	4.04*	.40	3.75*	.55	−1.79*	−.54	2.61*	.61
4	3.18*	.14	3.20*	.22	−1.77*	−.17	2.30*	.19
5	3.56*	.09	3.57*	.13	−1.78*	−.14	2.40*	.14
Declining communities								
Model 1	1.17	.02	.78	.02	−.45	−.01	.85	.02
2	1.09	.01	.44	.01	−.80	−.02	1.52	.04
3	.73	.09	.27	.06	−.74	−.20	1.36	.33
4	.73	.01	.27	.01	−.74	−.02	1.30	.04
5	.74	.01	.27	.01	−.73	−.02	1.30	.03

* Significant at the .05 level.

Table 6. Metric-Slopes (b) and Standardized Slopes (Beta, in Parentheses) for Model 4 of Changes in the Administrative Proportion Associated with Three Log Population Change Measures

Independent variables and administrative group	Growing			Declining		
	Total	County seat	Not co. seat	Total	County seat	Not co. seat
Log Pop						
Administrative	3.18*	3.66*	2.09	.73	1.91	−.60
	(.14)	(.14)	(.11)	(.01)	(.03)	(−.01)
Professional	3.20*	2.65*	3.36*	.27	1.13	−.90
	(.22)	(.16)	(.28)	(.01)	(.03)	(−.02)
Managerial	−1.77*	−2.28*	−1.28*	−.74	−.54	−1.11
	(−.17)	(.19)	(−.14)	(−.02)	(−.02)	(−.04)
Clerical	2.30*	3.53*	1.47*	1.30	.79	2.12
	(.19)	(.24)	(.15)	(.04)	(.02)	(.06)
Log County Rural Pop						
Administrative	1.82*	2.31*	.71	2.33*	1.71	4.07
	(.09)	(.12)	(.03)	(.09)	(.08)	(.11)
Professional	1.21*	1.02	1.75	.31	.15	.18
	(.09)	(.08)	(.11)	(.02)	(.01)	(.01)
Managerial	−.24	.01	−.90	.50	−.02	1.49
	(−.02)	(.00)	(−.08)	(.03)	(.00)	(.08)
Clerical	1.27*	1.52*	.85	1.06	.44	2.97
	(.11)	(.14)	(.07)	(.07)	(.04)	(.14)
Log County Pop						
Administrative	2.87*	3.93*	1.44	2.08	1.72	3.74
	(.11)	(.14)	(.06)	(.05)	(.05)	(.09)
Professional	2.36*	2.21*	2.68*	.37	.74	−.12
	(.13)	(.12)	(.16)	(.02)	(.03)	(−.00)
Managerial	−1.13*	−.97	−1.35	.20	−.40	1.23
	(.09)	(−.08)	(−.11)	(.01)	(−.02)	(.06)
Clerical	2.48*	3.38*	1.31	1.15	.00	3.40
	(.16)	(.21)	(.10)	(.05)	(.00)	(.13)
(No. of Places)	(868)	(573)	(295)	(527)	(325)	(202)

NOTE: Beta coefficients here are also zero order correlations.
* Significant at the .05 level.

Though none of these differences between corresponding coefficients for county seats and non–county seats is significant, the results at least for the clerical category are consistent with expectations. The expansion of clerical activities is greater in growing county seats than other growing places because of the greater importance of administrative functions in county seats. Conversely, these functions should not decrease as much with population decline as in places that are not county seats.

Since our places were selected to emphasize their labor-force and shopping autonomy from metropolitan areas, they probably serve as trade, work, and social centers for towns in their hinterlands. Consequently, the changing population of the place per se might not be the only variable affecting growth and/or decline. Since the work of Galpin (1915), rural communities have been conceived of as trade and social centers of a wider area. Over time, these areas may have expanded and interpenetrated, much as have metropolitan areas (McKenzie, 1933; Hawley, 1950). Until recently, an important difference has been that organizational expansion has been accompanied by population concentration in nonmetropolitan areas, but deconcentration in metropolitan areas. There was peripheral population growth in the 1960s around larger nonmetropolitan places, however; and since 1970, nonmetropolitan rural areas in general have grown more rapidly than urban, paralleling the metropolitan trend of long standing (Fuguitt, Lichter, and Beale, 1981).

Drawing on the theory of ecological expansion, we hypothesized an association between change in administrative structure and both total and rural population change in the counties where our places are located. Under the head "Log county rural pop" of Table 6 we give results using change in the population of the counties, excluding places with population of 2,500 in 1960, as the residualized independent variable. Results for growing places are all smaller than when only the population of the place itself is considered. The overall administrative and clerical proportions are significantly associated with peripheral growth for county seats, as are the administrative, professional, and clerical proportions for all places. For declining places, only the overall administrative component is significantly associated with outlying rural population change.

Finally, the results for the association with total county population follow closely the patterns above, with smaller sized coefficients than for the Log Pop variables, fewer of which are significant. For declining places, none of the administrative proportions is significantly associated with change in county population.

In all, there is some evidence that change in total county population and outlying rural population is associated with change in the administrative ratios for growing centers, particularly county seats. These b-coefficients, however, with only two exceptions, are smaller than the corresponding coefficients for change in the place's population. Thus evidence for the expansion hypothesis is not strong, perhaps because expansion of the rural community in an organizational sense until very recently has not been accompanied by functionally integrated population deconcentration. Therefore, rural tributary areas may not be as dependent for routine functioning on county seats as many suburbs are on metropolitan cities.

SUMMARY AND DISCUSSION

Building upon the contributions of organizational and community theorists, we assessed how three components of community sustenance structure (managerial, clerical, and professional) respond to population size, growth, and decline. For relatively autonomous nonmetropolitan U.S. cities which grew between 1960 and 1970, our hypotheses for the professional and clerical components were confirmed, with positive associations between population size and growth and the proportion of employed persons in these occupational groups. For the managerial component, we found a negative association, consistent with most previous research using organizations as units. For these communities, however, the results for managers may be complicated by the proportion

of employed who are proprietors, particularly those in retail trade.[4]

The same associations were also generally found for declining places, but typically they were not statistically significant. However, since we are working with a universe of communities, we conclude that a structural symmetry may exist between decline and growth situations, but that administrative components respond less to decline than to growth. This interpretation is consistent with the notion of administrative "featherbedding" in organizational theory; that is, positions related to administration are more likely to be retained than might be expected.

We initially hypothesized that the association between population size or change and clerical functions would be the largest among the three administrative components. This tended to hold for the cross-sectional correlations, but not for all dynamic analyses. A suggestive difference emerged in the clerical function between county seats and other places, with county seats having a higher association with clerical functions than other places in the growing situation, but less association than other places in the declining situation. In all likelihood, clerical and record-keeping activities are more important in county seats growing in population, but remain more or less stable despite population decline.

We also associated change in managerial, clerical, and professional components of communities with population in the surrounding areas. Our analysis provided some evidence that places may respond to population change in adjacent or more inclusive functional areas such as rural or total county populations. This relationship is particularly strong for growing county seats. But our results also indicate that it may be premature to consider rural areas surrounding nonmetropolitan cities as analogous to the suburban rings of metropolitan central cities, where previous support for the ecological expansion theory has been found.

The Boulding formulation implies that population change leads to nonproportional structural change; however, it is not the only possible framework for interpretation. Population and organization are reciprocally interrelated, and parallel formulations for the demographic-ecological perspective often treat population change as a dependent variable. Internal and/or external organizational changes can also lead to an increased number and proportion of administrative workers in a community and thereby attract additional population. The issue here is the direction of causation between population and administrative structure. We therefore performed both Pelz and Andrew's (1964) cross-lag correlation analysis and Heise's (1970) two-variable, two-wave regression analysis. With one exception (growing non–county seats) all cross-lagged correlation analyses and two-variable, two-wave analyses gave consistent support for the direction from population to the administrative proportion. This issue of the primary direction of causality is not settled, and should be addressed in future research along with the consequences of dealing with entities that do not have complete autonomy, whatever the social units employed.[5]

In sum, although theory and research on the implications of population size, growth, and decline for morphological structure have focused on the full spectrum of social units ranging from small groups (Simmel, 1902–1903) to nation states (Mayhew, 1973; Nolan, 1979), most hypotheses, while specified in dynamic terms, have been assessed with cross-sectional analyses. Whereas our dynamic findings did not replicate the static ones precisely, they were consistent. Consequently, we have greater con-

[4] The 1980 Census will make it possible to examine the relationship between population change and change in the proportion of retail proprietors, separately from others in the managerial category. This new data may facilitate study of the somewhat different predictions provided by organization theory and central-place theory for change in retail activities of nonmetropolitan urban places (see Berry, 1967).

[5] Haire (1959) considered the issue of autonomy by differentiating and associating the growth of the internal and external functions of organizations. The parallel with the basic-nonbasic distinction for community economic activities immediately comes to mind, although the difficulties in actually dividing activities in this way, particularly using census sources, are well known.

fidence in the results of previous cross-sectional studies which related population size and community structure to underlying dynamic processes. Moreover, in both the static and dynamic analyses, we presented a variety of statistical methodologies which yielded quite consistent results. This consistency leads us to suggest that needed substantive work should not be sacrificed to excessive debate over alternative statistical methods.

REFERENCES

Bereiter, C.
1962 "Some persisting dilemmas in the measurement of change." Pp. 3–20 in C. W. Harris (ed.), Problems in Measuring Change. Madison: University of Wisconsin Press.

Berry, Brian J. L.
1967 Geography of Market Centers and Retail Distribution. Englewood Cliffs, N. J.: Prentice-Hall.

Blalock, H. M., Jr.
1979 "Measurement and conceptualization problems: the major obstacle to integrating theory and research." American Sociological Review 44:881–94.

Blau, Peter M.
1970 "A formal theory of differentiation in organizations." American Sociological Review 35:201–18.

Blau, Peter M., Wolf V. Heydebrand, and Robert E. Stauffer
1966 "The structure of small bureaucracies." American Sociological Review 31:179–91.

Bohrnstedt, George W.
1969 "Observations on the measurement of change." Pp. 113–34 in Edgar F. Borgatta and George W. Bohrnstedt (eds.), Sociological Methodology 1969. San Francisco: Jossey-Bass.

Boulding, Kenneth
1953 "Toward a general theory of growth." The Canadian Journal of Economics and Political Science 19: 326–40.

Christaller, Walter
1966 Central Places in Southern Germany. (Tr. by Carlisle W. Baskin.) Englewood Cliffs, N. J.: Prentice-Hall.

Cronbach, Lee J. and Lita Furby
1970 "How we should measure change—or should we?" Psychological Bulletin 74:68–80.

Duncan, Otis Dudley and Albert J. Reiss, Jr.
1956 Social Characteristics of Urban and Rural Communities, 1950. New York: Wiley.

Duncan, Otis Dudley and Leo F. Schnore
1959 "Cultural, behavioral, and ecological perspectives in the study of social organization." American Journal of Sociology 65: 132–46.

Durkheim, Emile
1893 De la Division du Travail Social. Paris: Alcan. Translated by George Simpson as The Division of Labor in Society. New York: Macmillan, 1933.

Erbe, William
1977 "System size and the administrative component in small urban places." Madison, Wisconsin: University of Wisconsin Center for Demography and Ecology Working Paper 77–22.

Freeman, John Henry and Michael T. Hannan
1975 "Growth and decline processes in organizations." American Sociological Review 40:215–28.

Frisbie, W. Parker
1980 "Theory and research in urban ecology: persistent problems and current progress." Pp. 203–19 in H. M. Blalock, Jr. (ed.) Sociological Theory and Research: A Critical Appraisal. New York: Free Press.

Frisbie, W. Parker and Dudley L. Poston, Jr.
1975 "Components of sustenance organization and nonmetropolitan population change; a human ecological investigation." American Sociological Review 40:773–84.
1978 Sustenance Organization and Migration in Nonmetropolitan America. Iowa City: University of Iowa Community Research Center.

Fuguitt, Glenn V.
1965 "County seat status as a factor in small town growth and decline." Social Forces 44:245–51.

Fuguitt, Glenn V. and Calvin L. Beale
1977 "Recent trends in city population (growth and distribution." Pp. 13–27 in Herrington J. Brice (ed.), Small Cities in Transition: The Dynamics of Growth and Decline. Cambridge, Mass.: Ballinger.

Fuguitt, Glenn V. and Donald R. Field
1972 "Some population characteristics of villages differentiated by size, location and growth." Demography 9:295–308.

Fuguitt, Glenn V., Daniel T. Lichter, and Calvin Beale
1981 Population Deconcentration in Metropolitan and Nonmetropolitan Areas of the United States, 1950–1975. Madison: University of Wisconsin–Madison Applied Population Laboratory Population Series 70–15.

Fuguitt, Glenn and Stanley Lieberson
1974 "Correlation of ratios or difference scores having common terms." Pp. 128–44, in Herbert L. Costner (ed.) Sociological Methodology 1973–74. San Francisco: Jossey-Bass.

Galpin, C. J.
1915 The Social Anatomy of an Agricultural Community. Madison, Wisconsin: University of Wisconsin Agricultural Experiment Station Research Bulletin 34.

Gibbs, Jack and Walter Martin
1959 "Toward a theoretical system of human ecology." Pacific Sociological Review 2:29–36.

Haire, M.
1959 "Biological models and empirical histories of the growth of organizations." Pp. 272–306 in M. Haire (ed.), Modern Organization Theory. New York: Wiley.

Halbwachs, Maurice
1938 Morphologie Sociale. Paris: Librarie Armand Colin. English translation by Otis Dudley Duncan and Harold W. Pfantz, Population and Society: Introduction to Social Morphology. Glencoe, Illinois: The Free Press, 1960.

Hannan, Michael T. and Nancy Branden Tuma
1979 "Methods for temporal analysis." Pp. 303–28 in A. Inkeles, J. Coleman, and R. Turner (eds.), Annual Review of Sociology. Palo Alto: Annual Reviews, Inc.

Hathaway, Dale E., J. Allan Beegle, and W. Keith Bryant
1968 People of Rural America. (A 1960 Census Monograph). Washington, D.C.: U.S. Government Printing Office.

Hawley, Amos H.
1950 Human Ecology: A Theory of Community Structure. New York: Ronald Press.
1968 "Human ecology." Pp. 323–32 in D. L. Sills (ed.), International Encyclopedia of the Social Sciences. New York: Macmillan.

Heise, David R.
1970 "Causal inference from panel data." Pp. 3–27 in E. F. Borgatta and G. W. Bohrnstedt (eds.), Sociological Methodology, 1970. San Francisco: Jossey-Bass.

Hendershott, G. E. and T. F. James
1972 "Size and growth as determinants of administrative-production ratios in organization." American Sociological Review 37:149–53.

Hibbs, Douglas
1974 "Problems of statistical estimation and causal inference in dynamic time series models." Pp. 252–308 in Herbert Costner (ed.), Sociological Methodology 1973–74. San Francisco: Jossey-Bass.

Johnston, J.
1972 Econometric Methods, 2nd ed. New York: McGraw-Hill.

Kasarda, John D.
1972 "The theory of ecological expansion: an empirical test." Social Forces 51:165–75.
1974 "The structural implications of social system size: a three level analysis." American Sociological Review 39:19–28.

Kasarda, John D. and Patrick Nolan
1979 "Ratio measurement and theoretical inference in social research." Social Forces 58:212–27.

Lazarsfeld, P. F. and H. Menzel
1961 "On the relation between individual and group properties." In A. Etzioni (ed.), Complex Organizations. New York: Holt.

MacMillan, Alexander and Richard L. Daft
1979 "Administrative intensity and ratio variables: the case against definitional dependency." Social Forces 58:228–49.

Mayhew, Bruce H.
1973 "System size and ruling elites." American Sociological Review. 38:468–75.

Mayhew, Bruce H. and Roger Levinger
1976 "Size and the density of interaction in human aggregates." American Journal of Sociology 82:86–110.

McKenzie, R. D.
1933 The Metropolitan Community. New York: McGraw-Hill.

Nolan, Patrick D.
1979 "Size and administrative intensity in nations." American Sociological Review 44:110–25.

Ogburn, William F.
1937 Social Characteristics of Cities. Chicago: International City Managers Association.

Ostrom, Charles W.
1978 Time Series Analysis: Regression Techniques. Beverly Hills: Sage.

Pelz, Donald and Frank Andrews
1964 "Detecting causal priorities in panel study data." American Sociological Review 29:836–48.

Priebe, John A., Joan Heinkel, and Stanley Greene
1972 1970 Occupation and Industry Classification Systems in Terms of Their 1960 Occupation and Industry Elements. Washington, D.C.: U.S. Department of Commerce, Bureau of the Census technical paper 26.

Rushing, William A.
1966 "Organization size and administration: the problems of causal homogeneity and a heterogeneous category." Pacific Sociological Review 9:100–8.

Schuessler, Karl
1974 "Analysis of ratio variables: opportunities and pitfalls." American Journal of Sociology 80:379–96.

Simmel, Georg
1902–3 "The number of members as determining the sociological form of groups." I and II. American Journal of Sociology 8:1–46, 138–96.

Smith, Richard and Robert Weller
1977 "Growth and structure of the metropolitan community." Pp. 76–147 in Kent Schwirian (ed.), Contemporary Topics in Urban Sociology. Morristown, N.J.: General Learning Press.

Spencer, Herbert
1877 The Principles of Sociology, Vol. 1. New York: D. Appleton and Company.

Starbuck, William
1965 "Organizational growth and development." Pp. 451–533 in J. G. March (ed.), Handbook of Organizations. Chicago: Rand McNally.

Stinchcombe, Arthur L.
1959 "Bureaucratic and craft administration of production." Administrative Science Quarterly 4:168–87.

Svalastoga, K.
1974 The Social System. Denmark: Akademisk Forlag.

Tarver, James D. and Calvin L. Beale
1968 "Population trends of southern nonmetropolitan towns, 1950 to 1960." Rural Sociology 33:19–29.
Thompson, D'Arcy
1917 On Growth and Form. Cambridge, Mass.: Cambridge University Press.
Thompson, James D.
1969 Organizations in Action. New York: McGraw-Hill.
Tsouderos, John E.
1955 "Organizational change in terms of a number of selected variables." American Sociological Review 20:206–10.
Wilson, Franklin D.
1978 "The organizational components of expanding metropolitan systems." Pp. 133–56 in Karl E. Taeuber, Larry L. Bumpass, and James A. Sweet (eds.), Social Demography. New York: Academic Press.
Winsborough, Hal H.
1960 "Occupational composition and the urban hierarchy." American Sociological Review 25:894–7.

RESEARCH NOTES

EFFECTS OF UNITED STATES PRESIDENTIAL ELECTIONS ON SUICIDE AND OTHER CAUSES OF DEATH*

MYRON BOOR
University of Pittsburgh and Western Psychiatric Institute and Clinic

The proportion of suicides and the proportion of deaths by other causes that occurred during September and October of Presidential election years were found to be significantly smaller than the proportions of September-October deaths in adjacent nonelection years. These results suggest that increased societal integration fostered by impending Presidential elections decreased mortality (especially suicides) during the two months preceding the elections. November–December results suggest that mortality reduction during September and October of election years is more long-lasting for suicide than for other causes of death. The collective effects on mortality of Presidential elections and other factors that affect societal integration may be substantial.

Durkheim (1897/1951) hypothesized that the suicide rates of societies are inversely related to the social integration of those societies. Phillips and Feldman (1973) found that the proportion of deaths from all causes in the United States during September and October of the Presidential election years from 1904 to 1968 was smaller than the proportion of deaths during those months in adjacent nonelection years. They concluded that the increased societal integration fostered by impending Presidential elections reduces mortality from causes of death other than suicide.

The present study compared the mortality levels from suicide and from other causes of death in September and October of Presidential election years with the respective mortality levels in those months of adjacent nonelection years in order to investigate the relative effects of impending Presidential elections on deaths by suicide and other causes. Lower levels of suicide during September and October of election years would suggest that suicidal persons are integrated into society to the extent that its political events affect their suicidal behaviors. A lack of difference between September–October suicide levels in election and nonelection years would suggest that suicidal persons are detached from society to the extent that they are not as influenced by its political events as are persons who die from other causes (Phillips and Feldman, 1973).

Suicidal crises often are momentary and transient (Seiden, 1977) whereas other causes of death (primarily physical diseases and dysfunctions) are likely to be more chronic and persistent. Thus, suicidal deaths might be less likely than deaths by other causes to increase during November and December of election years following a decrease during September and October. The November–December mortality levels by suicide and by other causes of death in election and nonelection years thus were compared in order to investigate the relative permanence of decreased September–October mortality levels in election years. This study provides additional evidence regarding the effects of social participation and societal integration on suicide and on deaths by other causes.

* Direct all correspondence to Myron Boor, Department of Psychiatry, 9th Floor APC Building, Rhode Island Hospital, Providence, RI 02902.

This research was supported by NIMH Psychiatric Epidemiology Training Grant MH-T32-15169-03. The writer is now at Rhode Island Hospital and Brown University.

METHOD

Mortality levels for suicide and for deaths by other causes were determined by procedures adapted from Phillips and Feldman (1973). The total number of suicides (deaths by other causes) in the United States that occurred during September and October in the 14 election years for which the necessary data (U.S. Bureau of the Census, 1911–1936; U.S. Public Health Service, 1937–1973) were available (1912, 1916, and 1928 to 1972) was divided by the total number of suicides (deaths by other causes) that occurred during those years to determine the proportion of suicides (deaths by other causes) that occurred during September and October of election years. Analogous procedures were used to determine the proportion of suicides (deaths by other causes) in the 28 adjacent nonelection years that occurred during September and October and to determine November–December mortality levels for election and nonelection years. Monthly suicide statistics were not reported in the census data prior to 1909 or for the 1921 and 1923 adjacent nonelection years; thus it was not possible to include the election years prior to 1912 or the 1920 and 1924 election years. The denominators of the proportions for election years were reduced by 1/366 to equate the number of days in election and nonelection years.

RESULTS

The proportion of September–October suicides in the 14 election years (.1626) was significantly smaller than the proportion of suicides during these months in the 28 adjacent nonelection years (.1647, $z = 2.20$, $p < .03$). The proportions of September–October suicides were smaller in election years than in adjacent nonelections years for 9 of the 14 individual comparisons ($z = 1.07$, $p < .15$); there thus was a fairly consistent tendency for suicide levels to be reduced in the two months preceding the elections. Likewise, the proportion of September–October deaths by other causes in election years (.1549) was significantly smaller than the proportion of deaths during these months in the nonelection years (.1558, $z = 9.36$, $p < .001$). The difference between the number of September–October deaths in election and nonelection years was proportionally greater for suicides than for deaths by other causes ($\chi^2 = 7.89$, df = 1, $p < .005$).

The proportion of November–December suicides in election years (.1575) tended to be smaller than the proportion of suicides during these months in the nonelection years (.1588, $z = 1.48$, $p < .15$). However, the proportion of November–December deaths by other causes in election years (.1734) was significantly larger than the proportion of deaths during these months in the nonelection years (.1716, $z = 17.96$, $p < .001$).

Analyses of these data with November and December deaths excluded from September–October mortality proportions and with September and October deaths excluded from November–December proportions (i.e., with mortality levels based on the remaining ten months rather than on the entire year) produced relationships that are comparable in direction and level of significance to those reported; thus these results are not artifacts of the partial interdependence of September–October and November–December mortality levels.

DISCUSSION

These data suggest that suicidal persons at the time of their suicides are integrated into society to the extent that their propensities toward suicide are reduced by the increased societal integration fostered by sociopolitical events such as Presidential elections. Impending Presidential elections in fact had a greater proportional effect on deaths by suicide than on deaths by other causes, though deaths by other causes also were reduced significantly prior to elections, as was reported by Phillips and Feldman (1973).

These data provide no evidence that persons whose suicides were prevented in the two months prior to Presidential elections showed increased suicidal behaviors in the weeks immediately following the elections. However, these data do suggest

that persons whose deaths by other causes were prevented during September and October of election years may have increased levels of mortality in November and December of those years. Thus, Presidential elections and other events that increase social integration may produce more long-lasting reductions in suicides than in deaths from other causes, perhaps because suicidal crises are more transient than the physical diseases and dysfunctions that cause most other deaths. (Phillips and Feldman (1973:685) indicated in a footnote that they found no increase in deaths following Presidential elections, but they did not indicate the procedures that they used to reach that conclusion.)

As Phillips and Feldman (1973) noted, the effects of Presidential elections on mortality levels are relatively minor. Approximately 1.3% of the suicides and 0.6% of the deaths from other causes that were expected to occur prior to elections did not occur. However, the collective effects of the many social events and conditions that may affect social integration and subsequent mortality (especially suicides) may be substantial. These events include Yom Kippur and birthdays (Phillips and Feldman, 1973), wars (Stengel, 1965), and unemployment (Boor, 1980). This and other studies suggest the importance of increasing the social integration of societies and of encouraging social participation by individual members of those societies.

REFERENCES

Boor, Myron
1980 "Relationships between unemployment rates and suicide rates in eight countries, 1962–1976." Psychological Reports 47:1095–101.

Durkheim, Emile
[1897]
1951 Suicide. New York: Free Press.

Phillips, David P. and Kenneth A. Feldman
1973 "A dip in deaths before ceremonial occasions: some new relationships between social integration and mortality." American Sociological Review 38:678–96.

Seiden, Richard H.
1977 "Suicide prevention: a public health/public policy approach." Omega 8:267–76.

Stengel, Erwin
1965 Suicide and Attempted Suicide. Bristol: Maggibbon & Lee.

U.S. Bureau of the Census
1911–1936 Mortality Statistics. Washington: Government Printing Office.

U.S. Public Health Service
1937–1973 Vital Statistics of the United States. Washington: Government Printing Office.

MEASURE FOR MEASURE: ON THE USE OF ENDOGENOUS ABILITY DATA IN SCHOOL-PROCESS RESEARCH*

KARL L. ALEXANDER
AARON M. PALLAS
Johns Hopkins University
MARTHA A. COOK
Case-Western Reserve University

The appropriate use of ability test performance in evaluating schools has been a subject of considerable debate. The present study evaluates three strategies for controlling input-level ability differences in such research: longitudinal, predetermined ability measures; cross-sectional, endogenous ability measures; and SES measures as proxies for ability. In general, endogenous measures of ability behave much like predetermined measures, while SES measures are, for many outcomes, seriously deficient as proxies for ability measures. We conclude that research that has relied upon cross-sectional data to control for student ability is not seriously flawed because of this, and that the omission of ability controls altogether is a much more serious failing.

"Input-output" studies of school effects and effectiveness (Astin, 1970; Averch et al., 1972; Herriott and Muse, 1973) seek to isolate the influences of schools and schooling by adjusting for academic resources that students already might possess upon entering a particular educational environment. Such resources include preschooling family influences, genetic endowments, relevant aspects of social background, and community characteristics. This perspective is exemplified in the school-process model, used extensively to estimate both social-psychological and school-organizational influences upon educational outcomes. Such models typically include as "inputs" to the schooling process measures of family SES and of students' academic competencies. Even when only in-school, cross-sectional test scores are available, such data are commonly used as though they reflected pre-existing differences in ability levels. That is, such data are used as input controls, rather than as endogenous outcomes which themselves might be influenced by the youth's personal history of academic experience. Practically all research in the status-attainment tradition uses testing data in this way.

In using endogenous ability data as though they were exogenous, sociologists implicitly assume that academic aptitude is largely unaffected by variations in school quality and educational experience. Neglecting such influences when they actually are non-negligible could produce substantial bias in the estimations from such models: in general, ability effects would tend to be overstated, and the potency of other factors which in fact influence endogenous test performance would be attenuated.

Conventional usage notwithstanding, the proper location of student ability in school attainment models actually has been the subject of much controversy over the years (see, in chronological order, Sewell and Armer, 1966, versus Boyle, 1966, Michael, 1966, and Turner, 1966; Bidwell and Kasarda, 1975 and 1976, versus Alexander and Griffin, 1976a and 1976b; Crain and Mahard, 1978, versus Eckland, 1979). Two issues arise recurrently in these exchanges: first,

---

* Direct all correspondence to Karl L. Alexander, Department of Social Relations, The Johns Hopkins University, Baltimore, MD 21218.

This research was funded through Grant No. NIE-G-80-0113 from the National Institute of Education to the Johns Hopkins University Center for the Social Organization of Schools. The conclusions and opinions expressed herein are solely the authors' and are not necessarily those of the NIE. We are grateful to Laura Salganik for her constructive comments on an earlier version of this paper.

whether cross-sectional measures of student ability reasonably can be used as proxies for preschool measures; and, second, whether other information on student family background is adequate to control for ability differences when testing data are unavailable (see also Wolfle, 1980).[1] These issues are not simply obscure details of model specification in the educational attainment literature, but involve fundamental questions regarding the responsiveness of tested abilities to variations in school quality and school experiences. Sustaining conventional practice implies trivial school influences upon student competencies; impugning it requires evidence for important, extrafamilial environmental and experiential influences upon tested academic aptitude.[2] To date, there is little concrete evidence to inform the debate.

The present study considers whether school-process analyses are appreciably distorted when endogenous testing data are used as though they were exogenous. We already have given some consideration to these issues as a subsidiary concern in an investigation of school effects on standardized test performance (Alexander, McPartland, and Cook, 1980). Focusing on SCAT (School and College Ability Tests) and STEP (Sequential Tests of Educational Progress) performance levels (see ETS, 1957a,b) as outcomes, we found that estimates of school effects and of high-school curriculum effects were substantially upwardly biased when family-background measures were used as proxies for academic ability.

In that study we also considered the consequences of using cross-sectional "ability" data (i.e., SCAT performance) to adjust for student competencies rather than assessments made prior to the student's entering a particular school setting.[3] In those analyses, which used performance on the STEP battery as the achievement criterion, school effects were substantially attenuated. That is, cross-sectional data overcontrolled for preexisting ability differences, resulting in marked underestimates of school influences. It turned out, then, that neither SES data nor cross-sectional testing data were suitable proxies for temporally prior ability measures: SES proxies led to overestimates of school influence, while cross-sectional test controls resulted in marked underestimates.

These are quite important conclusions. Unfortunately, this study was very narrow in its focus. It considered only school and curriculum effects on standardized test performance. We therefore do not know whether the use of ability proxies or cross-sectional ability measures is similarly problematic for other outcomes and for other sources of influence. Much of

[1] There does exist a modest literature, mainly in economics, which indicates that estimates of the financial returns to investments in schooling are substantially upwardly biased in the absence of suitable controls for pre-existing, economically relevant "resources" (i.e., academic ability and SES background). See, for example, Chamberlain, 1977; Griffin, 1976; Olneck, 1976; but also Griliches, 1977, for a somewhat different perspective. These studies generally are quite compatible with Wolfle's argument and evidence that SES measures are not adequate proxies for measures of academic competency in analyses that purport to estimate the influence of educational attainment on later outcomes. The same reasoning would make one wary of their adequacy in studies which estimate school-based influences upon educational attainments.

[2] Our framework implicitly adopts several important premises regarding such standardized tests and the character of the qualities that they tap. The aspects of cognitive functioning measured by such tests are presumed to be important resources for success in school and for success in socioeconomic strivings generally in complex industrial society. Moreover, the instruments are presumed to assess such qualities with acceptable levels of reliability and validity. Unless stated otherwise, the framework itself is agnostic as to whether such qualities are predominantly inherited as opposed to developed, whether the tests are "culturally biased," and whether the qualities themselves are best thought of as "abilities," "aptitudes," or "achievements." We recognize that many would be unwilling to grant even the minimum assumptions mentioned above.

[3] In this study we also use SCAT to control for ability differences among students. Although the SCAT battery is not an IQ test, it is like many of the standardized test batteries used in school-process research. The ETS technical documentation describing the SCAT battery leaves the impression that it is intended to tap aptitude more than achievement, and at least one authority on the psychometrics of test construction has criticized the use of SCAT as an achievement outcome in Coleman et al.'s (1966) school-effects research because it actually indexes aptitude inputs (Kendrick, 1967:8).

the school-process literature pursues social-psychological issues, involving, for example, social supports and interpersonal influences as factors in school performance and in students' orientations toward their vocational and educational futures. Such subjective and interpersonal variables differ from testing outcomes in many important respects, and there can be little confidence that conclusions which apply to one will generalize to the other.

The present analysis, then, broadens the scope of our earlier study. Our basic strategy will be to compare three versions of simple school-process models for outcomes at several grade levels: 7, 9, 11, and 12. Within grade levels, alternative estimations will differ only in the strategy used to control for "input" differences in student ability levels: we will use SES measures as proxies for ability, an "in-school," usually cross-sectional, ability control, and a temporally prior ability measurement.[4] Somewhat different models will be used at each grade level.

The Seventh Grade. As seventh grade outcomes we consider the student's plans for college, his/her intended high-school curriculum, and performance on the STEP battery. These will be predicted first using background characteristics (race, sex, and parental education) and, as the measure of academic aptitude, fifth grade SCAT performance. Since our sample will be restricted to schools in which the seventh grade is a junior high school and the fifth grade is an elementary school, these ability controls not only are temporally prior to the outcomes, but they predate as well the level of schooling at which the outcomes are referenced. Hence, they are unambiguously predetermined relative to experiences and influences originating in that particular school context. These results, then, serve as our benchmark, and we compare them with those obtained when, first, no ability controls are used and, second, when seventh grade (i.e., cross-sectional) SCAT measurements are substituted for the fifth grade test scores.

The Ninth Grade. For this sample of schools, the ninth grade also is a junior high-school grade, so our ninth grade analysis again uses fifth grade SCAT data as the longitudinal ability control, this time contrasting it with its ninth grade counterpart. We use as dependent variables measures of curriculum plans, educational plans, STEP scores, and three indicators of significant others' influence (mother's encouragement for college, father's encouragement for college, and the college plans of the student's friends).

The Eleventh Grade. As ability controls for outcomes at the eleventh grade, we contrast eleventh and ninth grade SCAT assessments. Ninth grade SCAT scores precede high-school entry and hence are predetermined relative to all eleventh grade outcomes.

For this analysis we use a cross-sectional school-process model very much akin to those introduced by Sewell and his colleagues (Sewell and Hauser, 1980). It is a simple recursive specification in which STEP performance is the "ultimate" outcome. It is preceded by college plans, which in turn are preceded by curriculum enrollment, which is preceded by the three measures of significant others' influence. The first stage of the model consists of the various background influences that are used throughout these exercises: race, sex, and socioeconomic origins.

Although this organization of the eleventh grade data admittedly is rather arbitrary, it is at least as defensible as those that have been used in substantive research on these topics. The procedural questions that motivate this inquiry make some such heuristic specification desirable, since this allows us to consider how estimations for intervening influences on school process are affected by the choice of ability controls.

The Twelfth Grade. As twelfth grade

[4] We make no attempt to identify the specific sources of bias in cross-sectional data nor do we explore possibilities for better approximating longitudinal results cross-sectionally through various analytic means (e.g., unmeasured variables analysis, experimenting with assumptions regarding error structure, etc.). Boardman and Murnane (1980) provide a useful formal treatment of many of these latter issues. They begin with a "true" developmental model of educational achievement and demonstrate algebraically the respects in which the estimations obtained from several "degenerate" designs (e.g., cross-sectional, pre- and post-measures) depart from the model's true parameters.

outcomes, we consider, again, educational plans and the following additional outcomes: whether the student has yet to be admitted to a college, senior class rank, and SAT verbal and quantitative performance. The twelfth grade analysis fits in least well with our general intent and design, since we lack a cross-sectional SCAT assessment for this grade level. Instead, we use the eleventh grade measure as an in-school assessment and contrast it with that from the ninth grade. Despite this slippage, there are two reasons for reporting the twelfth grade results: some of the more interesting outcomes are available only at the twelfth grade (SATs, class rank, and college admission), and using eleventh grade school-process variables as predictors of twelfth grade outcomes affords a second opportunity to see how the choice of ability controls affects the estimated influence of intervening variables. Also, the specification of the eleventh grade data as antecedent to the twelfth grade outcomes is more secure than is the arbitrary ordering of the eleventh grade variables for the eleventh grade model. It thus seems to us a sensible and informative safeguard to include the twelfth grade outcomes in this exercise.

Should our results be reasonably similar throughout these analyses, then it would seem that apprehensions regarding the use of endogenous ability data as though they were predetermined in school-effects and school-process research are misplaced. On the other hand, if the disparities across alternative assessments are at all marked, we may be able to determine if they are especially severe at some particular stages of schooling, for certain classes of predictors, or for certain kinds of outcomes. If the biases are conveniently patterned along some of these lines, we will have some guidance as to how to improve practice in the future and as to which claims from the literature are most suspect.

METHOD

Sample. The data for these analyses were collected between 1961 and 1969 for the Study of Academic Prediction and Growth conducted by the Educational Testing Service. Seventeen communities in the United States participated in the Growth Study. The communities were purposively selected to vary on geographic location, school system size, and proportion of the community's high-school graduates continuing on to higher education (see Hilton, 1971, for more detail on the communities chosen). The initial Growth Study student sample, contacted in 1961, consisted of eleventh and twelfth graders in 27 high schools in these 17 communities, plus fifth, seventh, and ninth graders in their elementary and junior high feeder schools. These grade levels subsequently were resurveyed at least every two years during the period 1961 to 1969. We use data from the cohort first surveyed as fifth graders in 1961.

In each analysis, we limit the samples to individuals for whom there is complete information on all pertinent variables. Thus, at each grade level, all of the regression equations are based on the same cases. As mentioned earlier, we have restricted our sample to schools in which the fifth grade is in an elementary school, the seventh and ninth grades are in a junior high school, and the eleventh and twelfth grades are in a senior high school.

Variables. Our data derive mainly from questionnaires administered in grades seven, nine, eleven, and twelve and from standardized tests administered to students in grades five, seven, nine, eleven, and twelve. Although the specific source of information on a particular variable often will differ across grade levels, the questions and instruments used were either identical or closely parallel across all phases of data acquisition. For the seventh, ninth, and eleventh grade analyses, all of the variables except the predetermined SCAT are measured at the grade level of the analysis. For the twelfth grade analysis, only the dependent variables are measured at the twelfth grade, and all independent variables are assessed at the eleventh grade (except for the predetermined SCAT measure which is from the ninth grade). We used the following variables and abbreviations:

1. Race, with blacks coded 1 and nonblacks 0 (from school records, counselor reports, and yearbooks); RACE.

2. Sex, with women coded 1 and men 0; SEX.

3. Father's Education, scaled in years of school completed; FAED.

4. Mother's Education, scaled in years of school completed; MOED.

5. Academic Ability, the School and College Ability Tests composite score (see ETS, 1957a). Reliabilities for the SCAT composite are .95, .96, .95, and .95 for the 5th, 7th, 9th, and 11th grades (ETS, 1957a); SCAT.

6. Academic Achievement, the Sequential Tests of Educational Progress composite score (see ETS, 1957b). Using the formula suggested by Nunnally (1967) and STEP subtest reliabilities (ETS, 1957b), we estimated the reliability of the STEP composite to be .98, .98, and .97 for the 7th, 9th, and 11th grades; STEP.

7. Curriculum Placement, student reports collapsed to a dichotomy with college preparatory coded 1; CURRIC.

8. Peers' Plans, the students' reports of their friends' college plans. The five options were coded as percentages intending to go to college: 10, 30, 50, 70, and 90; PEER PLANS.

9. Father's Encouragement for College, the students' reports of their fathers' support for their post-secondary schooling. The original five response options (ranging from "strongly favors it" to "strongly opposes it") were coded as percentages in the same metric as PEER PLANS; FATHER'S ENC.

10. Mother's Encouragement for College, same as above for father's; MOTHER'S ENC.

11. Educational Plans, coded as a dichotomy, with 1 signifying plans to go to college; ED PLANS.

12. Curriculum Plans, student reports collapsed to a dichotomy with college preparatory coded 1; CURRIC PLANS.

13. P/SAT Verbal Scores, available only in the senior year, are the verbal scores on the SAT when available, otherwise the PSAT (Preliminary Scholastic Aptitude Test) rescaled to the SAT metric; P/SAT-V.

14. P/SAT Quantitative Scores, available only in the senior year, are the quantitative scores on the SAT when available, otherwise the PSAT rescaled to the SAT metric; P/SAT-Q.

15. Acceptance Status, whether the student had been admitted to college at the time of the senior year survey, with "yes" coded 1 and "no" as 0; COLLEGE ADMISSION.

16. Senior Class Rank, obtained from school records and coded as the percentage of students below the respondent; CLASS RANK.

Means and standard deviations for all of the measures used in the analysis are presented in Table 1.

RESULTS

Our results are displayed in a series of tables, one for each of the four grade levels we consider. Table 2 presents regressions of eleventh grade outcomes conforming to the cross-sectional school process model outlined earlier. Since this sort of model is most commonplace in the literature, we consider it first.

The equations employing longitudinal ability controls (9th grade SCAT scores) and those employing cross-sectional ability controls (11th grade SCAT scores) are strikingly similar. This holds for both the regression estimations and the proportions of explained variance. These are Equations II and III for each outcome. The only notable differences are observed for STEP performance: when cross-sectional SCAT controls are used the model underestimates the effect of educational plans and overstates the effect of sex relative to the results obtained with ninth grade SCAT controls. The attenuation of the plans coefficient is consistent with the expectation that cross-sectional testing data would, if anything, tend to overcontrol for ability differences, and this is the tendency reflected in the other minor differences scattered elsewhere throughout the table. The upward bias in the sex coefficient, on the other hand, does not conform to this pattern. We suspect this has to do with the particular mix of quantitative and nonquantitative items in the ninth and eleventh grade SCAT batteries, but have not yet checked this thoroughly. Apart from these differences involving STEP performance, the use of cross-sectional ability data as proxies for measures obtained prior to high-school entry results in very little distortion.

Our second concern, the adequacy of SES measures as proxies for direct assessments of student ability levels, is addressed by comparing Equations I with Equations II or III in Table 2. For peers' college plans, curriculum location, and standardized achievement test performance, the omission of ability measures

Table 1. Means and Standard Deviations for Variables Used in Later Analyses

	7th grade (N=867)	9th grade (N=2119)	11th grade (N=1704)[a]	12th grade (N=1704)[a]
RACE	.12	.15	.12	.12
	(.32)	(.36)	(.32)	(.32)
SEX	.55	.52	.50	.50
	(.50)	(.50)	(.50)	(.50)
FAED	13.42	12.85	12.79	12.79
	(2.53)	(2.82)	(2.94)	(2.94)
MOED	13.20	12.69	12.57	12.57
	(2.32)	(2.41)	(2.40)	(2.40)
Predetermined				
SCAT[b]	258.38	257.33	284.99	284.99
	(8.79)	(8.07)	(11.66)	(11.66)
Endogenous				
SCAT	272.55	284.73	290.92	290.92
	(11.62)	(12.78)	(13.26)	(13.26)
STEP	1620.36	1673.83	1722.34	
	(83.46)	(84.45)	(77.41)	
PEER PLANS		62.16	62.58	62.58
		(28.46)	(26.69)	(26.69)
FATHER'S ENC		81.19	83.83	83.83
		(19.76)	(13.38)	(13.38)
MOTHER'S ENC		82.66	84.41	84.41
		(17.75)	(12.70)	(12.70)
CURRIC			.62	.62
			(.49)	(.49)
ED PLANS	.68	.73	.76	.68
	(.47)	(.45)	(.43)	(.47)
CURRIC PLANS	.45	.66		
	(.50)	(.48)		
P/SAT-V				409.43
				(119.50)
P/SAT-Q				428.80
				(130.22)
COLLEGE ADMISSION				.20
				(.40)
CLASS RANK				63.32
				(28.63)

NOTE: Standard deviations are in parentheses.

[a] 11th and 12th grade analyses are based on the same cases.

[b] Measured in the 5th grade for 7th grade and 9th grade analyses; measured in the 9th grade for the 11th grade and 12th grade analyses.

makes a considerable difference. For instance, in the equations lacking ability controls, being black has a significant negative impact on each of these outcomes. When a temporally prior measure of ability is included as an additional control, this pattern is substantially altered. Now being black has a small, albeit nonsignificant, positive effect on peers' plans and curriculum, and the effect of race on STEP performance, although still significantly negative, is greatly reduced. The equations for standardized test performance are of special interest because they include as predictors the full set of variables tapping in-school sources of influence on junior year educational outcomes. It is quite clear in these instances that inclusion of the proper, or for that matter any, ability measure substantially reduces estimates of the importance of in-school processes for determining educational achievement. The differences here generally are much larger than those observed when comparing results under alternative ability controls.

In other respects, though, the results are not much affected by the omission of ability. This is especially so for junior year educational plans and reports of parental educational encouragement. These variables themselves, moreover, are rather

Table 2. Alternative Estimates of an Eleventh Grade School Process Model Under Three Strategies for Controlling Ability (N = 1704)

Dependent Variable		9th GRADE SCAT	11th GRADE SCAT	RACE	SEX	FAED	MOED	PEER PLANS	FATHER'S ENC	MOTHER'S ENC	CURRIC	ED PLANS	R^2
		Independent Variables											
PEER PLANS	I			−5.134*	−2.487*	1.965*	2.140*						.146
				(−.062)	(−.047)	(.216)	(.192)						
	II	.689*		2.236	−2.249*	1.420*	1.625*						.217
		(.301)		(.027)	(−.042)	(.156)	(.146)						
	III		.625*	2.173	−1.321	1.351*	1.657*						.221
			(.311)	(.026)	(−.025)	(.149)	(.149)						
FATHER'S ENC	I			.089	−1.129	.877*	.368*						.058
				(.002)	(−.042)	(.193)	(.066)						
	II	.168*		1.882	−1.071	.744*	.242						.075
		(.146)		(.045)	(−.040)	(.164)	(.043)						
	III		.159*	1.950	−.832	.721*	.245						.077
			(.158)	(.047)	(−.031)	(.158)	(.044)						
MOTHER'S ENC	I			2.388*	−2.569*	.669*	.599*						.066
				(.061)	(−.101)	(.155)	(.113)						
	II	.133*		3.815*	−2.523*	.564*	.500*						.078
		(.122)		(.097)	(−.099)	(.130)	(.094)						
	III		.129*	3.897*	−2.328*	.542*	.499*						.080
			(.135)	(.099)	(−.092)	(.126)	(.094)						
CURRIC	I			−.124*	.030	.024*	.016*	.006*	.003*	.006*			.315
				(−.082)	(.031)	(.146)	(.077)	(.310)	(.074)	(.164)			
	II	.015*		.029	.030	.016*	.008	.004*	.002	.006*			.409
		(.364)		(.019)	(.031)	(.099)	(.041)	(.216)	(.052)	(.162)			
	III		.014*	.029	.050	.015*	.009	.004*	.002	.006*			.415
			(.378)	(.019)	(.052)	(.091)	(.045)	(.210)	(.050)	(.159)			
ED PLANS	I			.082*	.047*	.001	.009*	.003*	.002*	.006*	.255*		.326
				(.062)	(.055)	(.009)	(.052)	(.182)	(.067)	(.192)	(.290)		
	II	.002*		.104*	.047*	.001	.008	.003*	.002*	.007*	.234*		.328
		(.065)		(.078)	(.055)	(.004)	(.047)	(.172)	(.065)	(.195)	(.265)		
	III		.003*	.112*	.052*	.000	.008	.003*	.002*	.007*	.223*		.331
			(.095)	(.084)	(.061)	(.001)	(.047)	(.168)	(.064)	(.197)	(.253)		
11th GRADE STEP	I			−59.718*	1.078	1.786*	1.965*	.417*	.361*	−.490*	57.640*	15.802*	.415
				(−.248)	(.007)	(.068)	(.061)	(.144)	(.062)	(−.080)	(.362)	(.087)	
	II	4.760*		−16.249*	2.705	.370	.411	.143*	.243*	−.194	16.413*	8.178*	.730
		(.717)		(−.068)	(.017)	(0.14)	(.013)	(.049)	(.042)	(−.032)	(.103)	(.045)	
	III		4.752*	−12.493*	10.005*	−.111	.673	.112*	.220*	−.150	10.393*	3.334	.811
			(.814)	(−.052)	(.065)	(−.004)	(.021)	(.039)	(.038)	(−.025)	(.065)	(.018)	

NOTE: Standardized coefficients are in parentheses. Asterisks indicate coefficient is equal to or greater than twice its standard error.

unresponsive to student ability (in both instances, the ability effects are positive and significant, but not large). It is partly because these net ability effects are so modest that its inclusion has so little bearing on the estimates of the other sources of influence.[5]

Regressions of senior year outcomes appear in Table 3. Once again, the alternative assessments using first ninth grade and then eleventh grade ability controls are quite similar; again, the use of an endogenous measure of ability as though it were predetermined does not appear to be especially problematic. Since the ordering here of the intervening and final endogenous outcomes is much more secure than was the case in the eleventh grade specification, these results provide valuable confirmation of the pattern discussed above.

Again too, omission altogether of an ability measure leads to large errors in assessing the importance of both background and school variables for some outcomes, especially SAT performance and class rank. As examples, consider the effects of race—a background variable—and curriculum location—a school-based variable—when estimated with and without a measure of ability in the prediction equations. In the latter instance the effects are quite substantial; in the former they all are much smaller, most to the point of being quite trivial.

For educational plans and reports of admission to college, in contrast, omission of the ability measure has minor consequences. Once again this pattern is observed for outcomes for which ability itself evidences very small, albeit statistically significant, net effects.

Overall, then, the pattern observed for twelfth grade outcomes parallels that at the eleventh grade. The coefficients obtained when cross-sectional controls for student ability are employed are practically interchangeable with those obtained when the conceptually preferred pre-entry measures are used. We thus find little evidence that serious bias results from use of endogenous ability data as though they were predetermined in models of this sort. The picture regarding the adequacy of SES controls as proxies for ability measures is not so clear. In some instances results are substantially distorted by this practice; in others the consequences are quite minor. It appears, moreover, that the problem is least severe for social-psychological outcomes (i.e., college plans and interpersonal supports) and most so for performance measures such as school achievement.

The 7th grade outcomes, which are presented in Table 4, are of particular interest because they are measured at the earliest grade for which we have a predetermined measure of ability. For the equations predicting curriculum plans and educational plans, the endogenous and predetermined ability measures behave in like fashion. The interchangeability of cross-sectional ability and predetermined ability is less assured, however, for 7th grade STEP performance. Here use of the cross-sectional measure attenuates race and SES (father's education and mother's education) affects and inflates the importance of sex. These are the largest differences observed thus far for the two SCAT controls.

The regressions for our 9th grade outcomes are reported in Table 5. For these outcomes, the effects of background variables are quite similar regardless of which ability control is used, much as we have found at other grade levels. Performance on the STEP battery is again something of an exception, however. When cross-sectional ability data are used, the coefficients for race and SES are sharply deflated and the coefficient for sex is slightly inflated relative to the equation containing the predetermined ability score, which of course serves as our benchmark.

Regarding SES measures as proxies for ability, we find relatively little distortion for most 7th and 9th grade outcomes. The exception is STEP performance. When

[5] To say that the omission of ability controls matters little in those instances in which ability itself matters little may seem trite, but it is precisely the opposite pattern that accounts for the substantial biases discussed in the paragraph above. The bias being revealed in comparing Equations I with Equations II or III reflects spuriousness in the first estimations resulting from the failure to control for a causally relevant, causally antecedent source of influence.

Table 3. Alternative Estimates of a Twelfth Grade School Process Model Under Three Strategies for Controlling Ability (N = 1704)

Dependent Variables		Independent Variables: 9th GRADE SCAT	11th GRADE SCAT	RACE	SEX	FAED	MOED	PEER PLANS	FATHER'S ENC	MOTHER'S ENC	CURRIC	11th GRADE ED PLANS	R^2
P/SAT-V	I			−50.738*	8.388	4.414*	3.239*	.565*	.340	−.513	82.429*	24.358*	.339
				(−.137)	(.035)	(.109)	(.065)	(.126)	(.038)	(−.055)	(.335)	(.087)	
	II	7.340*		16.292*	10.897*	2.230*	.842	.143	.159	−.057	18.856*	12.598*	.653
		(.716)		(.044)	(.046)	(.055)	(.017)	(.032)	(.018)	(−.006)	(.077)	(.045)	
	III		6.739*	16.232*	21.048*	1.724*	1.407	.133	.142	−.032	15.428*	6.676	.673
			(.748)	(.044)	(.088)	(.042)	(.028)	(.030)	(.016)	(−.003)	(.063)	(.024)	
P/SAT-Q	I			−68.067*	−39.343*	3.465*	2.183	.634*	.320	−.465	95.288*	36.617*	.391
				(−.168)	(−.151)	(.078)	(.040)	(.130)	(.033)	(−.045)	(.356)	(.120)	
	II	6.962*		−4.491	−36.964*	1.393	−.091	.234*	.148	−.032	34.991*	25.462*	.629
		(.623)		(−.011)	(−.142)	(.031)	(−.022)	(.048)	(.015)	(−.003)	(.131)	(.084)	
	III		6.329*	−5.170	−27.453*	.938	.462	.228*	.134	−.012	32.362*	20.010*	.640
			(.644)	(−.013)	(−.105)	(.021)	(.009)	(.047)	(.014)	(−.001)	(.121)	(.066)	
CLASS RANK	I			−6.401*	5.278*	−.108	.285	.112*	.150*	−.209*	6.900*	3.433	.070
				(−.072)	(.092)	(−.011)	(.024)	(.105)	(.070)	(−.093)	(.117)	(.051)	
	II	.862*		1.472	5.573*	−.365	.003	.063*	.129	−.156*	−.567	2.052	.145
		(.351)		(.017)	(.097)	(−.037)	(.000)	(.058)	(.060)	(−.069)	(−.010)	(.031)	
	III		.803*	1.579	6.786*	−.429	.066	.061*	.127	−.152*	−1.083	1.327	.152
			(.372)	(.018)	(.119)	(−.044)	(.006)	(.056)	(.059)	(−.067)	(−.018)	(.020)	
12th GRADE ED PLANS	I			.094*	−.015	.012*	.009	.003*	.003*	.001	.303*	.216*	.420
				(.065)	(−.016)	(.076)	(.045)	(.167)	(.084)	(.018)	(.317)	(.199)	
	II	.003*		.123*	−.013	.011*	.008	.003*	.003*	.001	.276*	.211*	.424
		(.080)		(.085)	(−.015)	(.070)	(.040)	(.156)	(.082)	(.023)	(.288)	(.194)	
	III		.003*	.126*	−.008	.011*	.008	.003*	.003*	.001	.271*	.207*	.425
			(.093)	(.087)	(−.009)	(.068)	(.040)	(.155)	(.081)	(.024)	(.283)	(.191)	
COLLEGE ADMISSION	I			−.109*	.041*	.005	−.005	.001*	.001	−.000	.129*	.059*	.096
				(−.088)	(.051)	(.035)	(−.033)	(.093)	(.045)	(−.001)	(.159)	(.064)	
	II	.005*		−.067*	.042*	.003	−.007	.001*	.001	.000	.090*	.052*	.107
		(.135)		(−.055)	(.053)	(.025)	(−.042)	(.075)	(.042)	(.008)	(.110)	(.056)	
	III		.005*	−.061*	.050*	.003	−.007	.001*	.001	.000	.081*	.047	.112
			(.161)	(−.050)	(.063)	(.021)	(−.041)	(.072)	(.041)	(.100)	(1.00)	(.051)	

NOTE: Standardized coefficients are in parentheses. Asterisks indicate coefficient is equal to or greater than twice its standard error.

Table 4. Alternative Estimations of a Seventh Grade School Process Model under Three Strategies for Controlling Ability (N = 867)

Dependent Variables		Independent Variables						
		5th GRADE SCAT	7th GRADE SCAT	RACE	SEX	FAED	MOED	R^2
CURRIC PLANS	I			−.135*	−.040	.020*	.016	.036
				(−.087)	(−.040)	(.102)	(.073)	
	II	.018*		.003	−.088*	.009	.010	.123
		(.325)		(.002)	(−.088)	(.045)	(.046)	
	III		.015*	.016	−.065*	.006	.007	.130
			(.341)	(.010)	(−.065)	(.032)	(.032)	
ED PLANS	I			−.084	.046	.030*	.026*	.078
				(−.058)	(.049)	(.163)	(.130)	
	II	.017*		.040	.003	.020*	.021*	.158
		(.312)		(.027)	(.003)	(.109)	(.103)	
	III		.016*	.076	.020	.016*	.017*	.198
			(.387)	(.052)	(.021)	(.084)	(.083)	
STEP	I			−77.443*	20.427*	7.587*	4.141*	.234
				(−.299)	(.122)	(.230)	(.115)	
	II	7.266*		−23.281*	1.670	3.181*	1.795*	.716
		(.765)		(−.090)	(.010)	(.096)	(.050)	
	III		6.074*	−14.932*	10.158*	1.884*	.472	.807
			(.846)	(−.058)	(.061)	(.057)	(.013)	

NOTE: Standardized coefficients are in parentheses. Asterisks indicate coefficient is equal to or greater than twice its standard error.

STEP performance is the dependent variable, the omission of ability as a control is critical, for in its absence the upward bias in the coefficients of the background variables is quite substantial.[6]

CONCLUSIONS

Our analyses indicate that, in general, cross-sectional, endogenous measures of student ability behave quite similarly to longitudinal, predetermined ability assessments, across a wide range of academic outcomes and for several different grade levels. An important exception to this pattern is performance on the STEP battery, especially at the earlier grade levels. We shall discuss this further below.

These analyses speak to several methodological issues. First, we conclude that endogenous measures of student ability in most instances are interchangeable with predetermined measures. This is a very important conclusion, in that much of what we have learned over the years regarding achievement processes, both in school and afterwards, derives from studies which might well have been found deficient in a critical aspect of model specification and research design. To know that the use of cross-sectional ability controls in this research does not seriously bias its conclusions makes that accumulated wisdom much more secure.[7] In this sense, our analysis sustains common practice and removes a nagging uncertainty.

Second, SES measures often are seriously deficient as proxies for ability, producing upwardly biased estimates of the effects of background and school-process variables across several grade levels and

[6] Following the suggestion of an anonymous reviewer, we also conducted these analyses using correlations corrected for the unreliability of the various standardized tests we employ. These test reliabilities are quite high, however, and the detailed results differed only minutely from the analyses we present. The overall patterning of results is not changed at all. The reliabilities we used are reported in the section of the text that discusses variable measurement. The reviewer also suggested adjusting for unreliability in the data reported by students. While this might be quite important were we interested in drawing substantive conclusions from these exercises, for our internal comparisons within grade levels the same student reports are used throughout. Hence, the differential reliability of such reports across grade levels would not affect our major conclusions.

[7] We must emphasize that this conclusion applies only to measures of standardized test performance. For other school-process variables, it is clear that in many instances cross-sectional data behave quite differently from longitudinal. See Cook and Alexander (1980) for further detail.

Table 5. Alternative Estimations of a Ninth Grade School Process Model under Three Strategies for Controlling Ability (N = 2119)

Dependent Variables		Independent Variables						
		5th GRADE SCAT	9th GRADE SCAT	RACE	SEX	FAED	MOED	R^2
PEER PLANS	I			−8.425*	−2.376*	2.054*	2.055*	.139
				(−.107)	(−.042)	(.203)	(.174)	
	II	.866*		−2.074	−3.179*	1.756*	1.662*	.188
		(.246)		(−.026)	(−.056)	(.174)	(.140)	
	III		.746*	1.778	−2.695	1.465*	1.440*	.220
			(.335)	(.023)	(−.047)	(.145)	(.122)	
FATHER'S ENC	I			−4.677*	.216	.708*	.934*	.049
				(−.086)	(.005)	(.101)	(.114)	
	II	.459*		−1.311	−.209	.550*	.725*	.078
		(.187)		(−.024)	(−.005)	(.078)	(.088)	
	III		.411*	.941	.041	.384*	.595*	.101
			(.266)	(.017)	(.001)	(.055)	(.072)	
MOTHER'S ENC	I			−3.980*	−.333	.482*	.861*	.041
				(−.081)	(−.009)	(.076)	(.117)	
	II	.376*		−1.220	−.682	.352*	.690*	.065
		(.171)		(−.025)	(−.019)	(.056)	(.093)	
	III		.358*	.918	−.486	.199	.566*	.089
			(.258)	(.019)	(−.014)	(.032)	(.077)	
CURRIC PLANS	I			−.171*	−.044*	.029*	.025*	.102
				(−.130)	(−.046)	(.171)	(.126)	
	II	.018*		−.037	−.061*	.022*	.017*	.180
		(.310)		(−.028)	(−.064)	(.133)	(.084)	
	III		.017*	.061*	−.051*	.015*	.011*	.253
			(.456)	(.046)	(−.054)	(.091)	(.056)	
ED PLANS	I			−.059*	.001	.027*	.027*	.089
				(−.048)	(.002)	(.172)	(.148)	
	II	.013*		.038	−.011	.023*	.021*	.135
		(.239)		(.031)	(−.012)	(.143)	(.115)	
	III		.012*	.110*	−.004	.017*	.017*	.180
			(.355)	(.089)	(−.004)	(.110)	(.093)	
STEP	I			−86.645*	13.159*	5.405*	5.163*	.264
				(−.371)	(.078)	(.180)	(.147)	
	II	7.274*		−33.299*	6.418*	2.900*	1.857*	.657
		(.695)		(−.142)	(.038)	(.097)	(.053)	
	III		5.800*	−7.274*	10.675*	.826*	.381	.824
			(.878)	(−.031)	(.063)	(.028)	(.011)	

NOTE: Standardized coefficients are in parentheses. Asterisks indicate coefficient is equal to or greater than twice its standard error.

numerous academic outcomes. The distortion is especially severe for achievement outcomes, such as STEP performance. However, for certain social-psychological outcomes, notably educational plans and interpersonal influences, the omission of ability appears not to be especially critical.

Third, the use of cross-sectional ability controls in equations predicting STEP performance tends to underestimate race and SES effects, especially in the early grade levels. These results are similar to the Alexander, McPartland, and Cook (1980) finding that the use of cross-sectional ability measures resulted in substantial underestimation of school effects in equations predicting STEP performance. From these comparisons, then, it seems that cross-sectional ability measures may sometimes overcontrol for background effects when academic achievement is the criterion variable, and this may be especially likely at earlier grade levels. This recognition balances somewhat our conclusion that failing to control for input-level ability differences may lead to exaggerated claims regarding the potency of both background and school process influences. If one is forced to choose between the lesser of two evils, however, in general the biases resulting

from the neglect of ability altogether are much more severe than those due to the use of cross-sectional controls when predetermined ones would be preferred conceptually.

Our results also have rather clear implications regarding yet another possibility for how cross-sectional testing data might be used. Several studies (Coleman, et al., 1966; Crain and Mahard, 1978) have evaluated such data as educational outcomes, with an eye toward identifying how school differences and educational experiences might bear upon them. These studies use only family and personal background characteristics as input controls. In these designs, then, prior performance levels are proxied by SES background, and standardized test performance is the focal dependent variable. The present analysis and our previous study (Alexander, McPartland, and Cook, 1980) identify this combination of control strategies and outcome variables as being subject to the most severe estimation biases. This, incidentally, is the analysis design used in Coleman's (1981) recent controversial and highly publicized comparison of the effectiveness of public and private high schools. Our exercises indicate that secure conclusions regarding differential school effectiveness cannot be drawn from such analyses. Whether Coleman is correct or not regarding the superiority of private schools must remain an open question; what is clear, however, is that his evidence on the issue is fundamentally flawed.

Finally, there is an important substantive implication that underlies these methodological conclusions. Our analyses provide indirect evidence that measured student aptitude is not very responsive to in-school processes and experiences. If it were, endogenous ability assessments presumably would evidence quite different properties from predetermined ones. Of course, our findings do not preclude the possibility of larger school effects at grade levels earlier than those considered in our analyses; however, in light of the present results, such effects must be demonstrated, and not merely assumed.

REFERENCES

Alexander, Karl L. and Larry J. Griffin
1976a "School district effects on academic achievement: a reconsideration." American Sociological Review 41:144–52.
1976b "Reply to Bidwell and Kasarda." American Sociological Review 41:755–62.

Alexander, Karl L., James M. McPartland, and Martha A. Cook
1980 "Using standardized test performance in school effects research." Pp. 1–33 in R. Corwin (ed.), Research in Sociology of Education and Socialization, Volume 2. Greenwich, Conn.: JAI Press.

Astin, Alexander W.
1970 "The methodology of research on college impact, I." Sociology of Education 43:223–54.

Averch, Harvey A., Stephen J. Carroll, Theodore S. Donaldson, Herbert J. Kiesling, and John Pincus
1972 How Effective is Schooling? A Critical Review and Synthesis of Research Findings. Santa Monica, Calif.: Rand.

Bidwell, Charles E. and John D. Kasarda
1975 "School district organization and student achievement." American Sociological Review 40:55–70.
1976 "Reply to Hannan, Freeman and Meyer, and Alexander and Griffin." American Sociological Review 41:152–60.

Boardman, Anthony E. and Richard J. Murnane
1980 "Using panel data to improve estimates of the determinants of educational achievement." Sociology of Education 53:113–21.

Boyle, Richard P.
1966 "On neighborhood context and college plans (III)." American Sociological Review 31:706–7.

Chamberlain, Gary
1977 "Education, income and ability revisited." Journal of Econometrics 5:241–57.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James M. McPartland, Alexander Mood, Frederic D. Weinfield, and Robert L. York
1966 Equality of Educational Opportunity. Washington, D.C.: U.S. Government Printing Office.

Coleman, James, Thomas Hoffer, and Sally Kilgore
1981 Public and Private Schools. Final Report to the National Center for Education Statistics, Contract NO. 300-78-0208. Chicago: National Opinion Research Center.

Cook, Martha A. and Karl L. Alexander
1980 "Design and substance in educational research: adolescent attainments, a case in point." Sociology of Education 53:187–202.

Crain, Robert L. and Rita E. Mahard
1978 "School racial composition and black college attendance and achievement test performance." Sociology of Education 51:81–101.

Eckland, Bruce K.
1979 "School racial composition and college at-

tendance revisited." Sociology of Education 52:122–5.

Educational Testing Service

1957a Cooperative School and College Ability Tests: Technical Report. Princeton: Cooperative Test Division, Educational Testing Service.

1957b Cooperative Sequential Tests of Educational Progress: Technical Report—Reading, Writing, Listening, Social Studies, Science, Mathematics. Princeton: Cooperative Test Division, Educational Testing Service.

Griffin, Larry J.

1976 "Specification biases in estimates of socioeconomic returns to schooling." Sociology of Education 49:121–39.

Griliches, Zvi

1977 "Estimating the returns to schooling: some econometric problems." Econometrica 45:1–22.

Herriott, Robert E. and Donald N. Muse

1973 "Methodological issues in the study of school effects." Pp. 209–36 in F. N. Kerlinger (ed.), Review of Research in Education, Volume 1. Itasca, Ill.: Peacock.

Hilton, Thomas L.

1971 A Study of Intellectual Growth and Vocational Development. Final Report (Project No. 6-1830; Grant No. OEG-1-6-061830-0650) to U.S. Department of Health, Education, and Welfare, Office of Education, Bureau of Research. Princeton: Educational Testing Service.

Kendrick, S. A.

1967 "The coming segregation of our selective colleges." College Review Board 66:6–13.

Michael, John A.

1966 "On neighborhood context and college plans (II)." American Sociological Review 31:702–6.

Nunnally, Jum C.

1967 Psychometric Theory. New York: McGraw-Hill.

Olneck, Michael R.

1976 "On the use of sibling data to estimate the effects of family background, cognitive skills and schooling." Discussion Paper No. 374–76. Madison, Wisconsin: Institute for Research on Poverty.

Sewell, William H. and J. Michael Armer

1966 "Neighborhood context and college plans." American Sociological Review 32:475–9.

Sewell, William H. and Robert M. Hauser

1980 "The Wisconsin longitudinal study of social and psychological factors in aspirations and achievements." Pp. 59–99 in A. C. Kerckhoff (ed.), Research in Sociology of Education and Socialization, Volume 1. Greenwich, Conn.: JAI Press.

Turner, Ralph H.

1966 "On neighborhood context and college plans (I)." American Sociological Review 31:698–702.

Wolfle, Lee M.

1980 "The enduring effects of education on verbal skills." Sociology of Education 53:104–14.

RELIGIOUS CONFORMITY IN AN AGE OF DISBELIEF: CONTEXTUAL EFFECTS OF TIME, DENOMINATION, AND FAMILY PROCESSES UPON CHURCH DECLINE AND APOSTASY*

HART M. NELSEN
Louisiana State University

Three contexts that affect religious socialization are discussed: denominational structure (church and sect vis-a-vis their environments), time (the religious climates of the 1950s and the present), and family processes (gender role and birth order). Data from 3,000 Minnesota youths aged 10–14 show that religious apostasy occurs more in the liberal, mainline Protestant denominations than among conservative bodies and that religious conformity is exhibited more by females and firstborn males. The interaction effects among gender, birth order, and denomination in predicting religiosity testify to the importance of considering contextual effects in religious socialization.

Individual behavior can be affected by both personal characteristics and social contexts. To examine the interaction between persons and environments requires measures of group as well as individual characteristics. Sociologists of religion can use denomination to reflect variation in social contexts of behavior. Studies focusing on the role of the family in religious socialization overlook the context in which socialization occurs. I will argue that normal family processes interact with religious organization and societal climate to transmit religiosity from one generation to the next. The roles of gender and birth order are important to the process.

Religious Organizations and the Environment

An especially critical time for religious organizations is the age when the children of members are to be converted or attracted to membership. At stake is the religious conformity of the youths. When the agents of socialization are successful, the patterns exhibited by the youths will resemble those of their forefathers. The transmission of religiosity across generations, however, is affected by the surrounding secular climate.

Some denominations are particularly vulnerable to secular forces; a critical distinction exists between churches, which basically accept the secular environment, and sects, which tend to reject it (Johnson, 1963:542). A church assumes an inclusive membership (Yinger, 1970: 259), while the sect is exclusive, having voluntary membership, in principle if not in fact, and demanding proof of commitment, discipline, or zeal. The church type of religious organization is especially subject to the "ebb and flow of sentiments" in society, because it has an inclusive membership and competes with nonreligious organizations with similar goals (Zald and Ash, 1966).

When the climate of society is favorable to religion, as in the 1950s (see Hoge, 1974), marked interdenominational differences in the training of youths and in their retention in churches should not be expected. In the 1960s and 1970s, however, the more churchlike, inclusive religious organizations—especially the mainline, more liberal bodies—evidenced loss of membership and attendance (Carroll et

---

* Direct all correspondence to Hart M. Nelsen, Department of Sociology, Louisiana State University, Baton Rouge, Louisiana 70803.

The collection and analysis of the Minnesota data reported in this article was supported by funds from the Boys Town Center for the Study of Youth Development at the Catholic University of America. The data subjected to secondary analysis and reported in the fifth footnote were from James Allen Davis (Principal Investigator), *General Social Surveys, 1972–1978: Cumulative Data* (National Opinion Research Center, producer, and the Roper Public Opinion Research Center, distributor). I owe special thanks to John D. McCarthy and Anne K. Nelsen for their insightful comments on drafts of this article.

al., 1979; Hoge and Roozen, 1979a, b). Bibby (1978:133–4) reports that conservative Protestants are better able than their liberal counterparts to socialize and retain their offspring. These findings (see also Wuthnow, 1976) suggest a cohort-historical effect—that today's youth, particularly those from liberal religious backgrounds, are not as receptive to organized religion as were youth in the past.[1]

In times when the culture does not affirm religion, competing value orientations are more pronounced. In describing "value structure theory," Hoge and Roozen (1979a:57–61) note that a shift toward a value not congruent with religion decreases religiosity. This theory posits that religious involvement is related to "conformity and tradition more than individual freedom and tolerance of diversity, social conservatism more than social change, and definite moral codes more than individualized moral orientations." Similarly, Mol (1971:304–6) notes that religion provides only one form of personal integration, which is linked with conservative sexual and political attitudes. Hoge and Roozen (1979b:328) write that the "broad cultural shift . . . has hit the affluent, educated, individualistic, culture-affirming denominations hardest. The shift occurred much more among the youth than the older adults" (see also Carroll et al., 1979:40).

Differences in Religiosity by Gender and Birth Order

Which youths should be especially susceptible to the influences of their parents? Past studies suggest that girls, regardless of ordinal position in the family, and firstborn boys exhibit conforming behavior, while later-born boys do not. Girls are more conformist than boys and, specifically, they are more religious (Argyle and Beit-Hallahmi, 1975:71–9; Putney and Middleton, 1961). Girls, who tend to have close contact with their mothers, are encouraged to be obedient and to exhibit responsibility, while boys are given greater autonomy (Rogers, 1969:263; Garai and Scheinfeld, 1968:243; Long and Henderson, 1973:198).

Sex differences in religious socialization are complicated by birth-order effects. Sampson and Hancock (1967) found that firstborn males conformed significantly more than secondborn, though there were no significant differences by birth order among females. Smith (1971:225) observes two theoretical interpretations: (1) the firstborn male has a closer relationship with his parents than would the later born, and, hence, "he might be expected to be more favorably oriented than his younger siblings toward his parents and toward their authority"; and (2) because of his age and development, the firstborn dominates those younger than he, obtaining a superordinate, conformist position, with the parents in fact often delegating some of their authority to him.[2] Smith's data showed that firstborn males were more likely than later-born males to accept parental authority; there were no birth-order differences among females or between only children and firstborn with siblings (Smith, 1971: 230).

Assuming that both firstborn and only children would be dependent on their parents, MacDonald (1969b:628; also see 1969a) hypothesized that they would be better socialized in religiosity than later borns. Employing data on college students raised in the fifties, MacDonald found that the later born were more likely to have no religious preference, that within the no-

[1] From an analysis of church attendance from 1939 through 1969, Wingrove and Alston (1974:33) concluded that "in addition to sex, specific cohort membership and the general societal environment [must] be taken into account as causative factors." Age becomes a salient variable when decline occurs. On patterns of church attendance by age see Wuthnow (1976), Roozen (1979:133; 1980:435), and Roof and Hoge (1980:412). Mol (1971:40) writes that "age makes a difference in denominations and countries where church attendance is at a low ebb, but . . . there are slight differences . . . where church attendance is very strong."

[2] The two interpretations can be contrasted for their implications for only children. In the first case, only children should also have a close relationship with their parents, and thus they should resemble or even exceed the firstborn in accepting their parents' authority. In the second situation, the only children, lacking siblings, could not find themselves in superordinate positions; thus they should be less likely than the first-born with siblings to accept parental authority.

preference category church attendance occurred more often among firstborn and only (especially female) children, and that among those with a religious preference there was no main effect for birth order. As with the lack of birth-order differences among females, who generally conform, few differences might be expected among church-related individuals growing up in a conservative era. In fact, MacDonald found birth-order effects only among the no-preference category.

DeBord (1969), who analyzed data from intact families with two children residing in Nashville in the late 1950s, reported (p. 564) that birth order itself did not predict adolescent church attendance. He specified sibling structure: when the two children were of the same sex, the older child attended more frequently; when they were of different sex, the girl exhibited a higher attendance level. While these findings are suggestive, they probably lack statistical significance; and given the conservative climate of the 1950s and the southern setting of the study, it is not surprising to find no striking birth-order effects.

Hypotheses

My two hypotheses are that (1) birth-order differences in religiosity should occur for Protestant males but not females, with later-born males showing the lowest levels of religiosity (comparing means by gender and birth order), and (2) birth-order differences should be pronounced among males in the mainline, liberal, Protestant denominations, while they should lack significance for those from more conservative backgrounds, especially Baptist and fundamentalist bodies.

Given that my data were collected in 1975, when the climate in the United States was not especially receptive to religion (on religious trends see Gallup, 1975), I expect birth order and gender to have an effect among youth in the more liberal religious bodies. Due to social control mechanisms and greater insulation against secular forces (Wilson, 1959), differences by gender and birth order are not expected among youth of the more conservative denominations. I expect Catholic youths to resemble those in the conservative Protestant denominations. Because my data are from only one point in time, the contextual impact of time cannot be determined.

I will also present data on the effect of maternal religiosity. Birth-order differences may exist among youths of the liberal, mainline denominations in part because the later born in these religious bodies might be more likely than the firstborn to identify their parents as not very religious, perhaps because these parents display diminishing levels of religiosity as their children grow older or their families expand in size. Then the means of the later born should become larger across the increasing levels of parental religiosity, approaching those of the firstborn.

The Data

The data were collected by questionnaire from 3,000 fourth through eighth graders from 23 public and Catholic schools in southern Minnesota. The data are weighted by parochial-public, residential, and grade-enrollment variables (resulting in the same N) because parochial school Catholics were oversampled. When weighted, the sample is judged to be representative of the larger, southern half of the state. Only the weighted means are reported here, but an analysis using the unweighted data did not change the findings.

An index of preadolescent religiosity was constructed of five items: frequencies of prayer, reading the Bible, and attending religious services; and beliefs concerning the Ten Commandments and biblical literalism. A factor analysis and an examination of the index reliability indicated that the composite index could be used. The responses to the five items (dichotomized close to 50–50) were summed. This index, as well as the sample and the weighting, is discussed more fully in Nelsen (1980).

Because there were insufficient numbers from the more liberal denominations (Episcopal, Presbyterian, and the United Church of Christ) to retain these as separate categories, they were combined. The liberal Methodists were analyzed sepa-

rately because their numbers were sizeable and because they should be most subject to secularizing forces. Mol (1976:82) writes that "church attendance can be used as a rough indication of the hold of a church or sect on its constituency;" taking the General Social Survey data (1972–1978) for the North Central part of the United States and comparing Methodists to other groups in church attendance, I found that Methodists were the least likely to attend. Following Stark and Glock (1968), I identified Lutherans as conservative and lumped them together because I felt that some of the respondents, given their young age (as low as 10), would be unable to identify their specific denomination. Thus, the denominational categories used in the analysis were (1) Methodist, (2) other liberal bodies, (3) Lutheran, and (4) Baptist and various fundamentalistic denominations that would be more sectlike. Catholics were excluded from the main part of the analysis, but summary data for them are presented where appropriate.

The questionnaire identified only children and permitted assigning ordinal position (firstborn, middle born, and last born). An examination of differences in religiosity by these three categories of birth order led to the conclusion that the real break comes between the firstborn and the later born. The literature on birth order emphasizes this dichotomy. Only children constitute a case apart from ordinal position and should not be lumped with the firstborn (Warren, 1966:48).

To encompass youth in broken homes as well as those in intact families, maternal rather than parental religiosity is employed. Maternal religiosity is measured by a two-item index that includes the mother's interest in religion and the frequency of her church attendance, as these are perceived by her child (on the utility of using the responses of youths here rather than maternal self-report, see Putney and Middleton, 1961; Acock and Bengtson, 1980; and Nelsen, 1980).

The Findings

Multiple Classification Analysis (MCA) was used in order to present the zero-order relationships (the magnitude of the relationship between each of the predictors and youth religiosity is summarized by eta, which is equivalent to r) and, for comparative purposes, to show the combined effect of the predictors assuming no interaction (see Table 1). MCA assumes no interaction among the predictors; the hypotheses, however, anticipate interaction. The R-squared from MCA is .418, while that from Analysis-of-Variance (ANOVA) is .468, which means that an additional 5% of the variance is explained with the inclusion of all interaction effects.

From examining the etas it can be seen that the best predictor of preadolescent religiosity is maternal religiosity, followed by Protestant denomination, gender, and

Table 1. Religiosity Means by Protestant Denomination, Maternal Religiosity, Gender, and Birth Order (MCA)

	Religiosity of Preadolescents	
Predictors	Unadjusted Means	Means Adjusted by the Other Predictors (MCA)
Protestants Denomination		
Methodist (N = 220)	1.68	1.89
Episcopal, Pres., U.C.C. (114)	2.67	2.43
Lutheran (791)	2.65	2.71
Baptist and Fundamentalist (165)	3.62	3.19
eta, beta	.32	.23
Maternal Religiosity		
Low (597)	1.67	1.74
Moderate (314)	2.84	2.81
High (380)	3.90	3.81
eta, beta	.58	.54
Gender		
Male (628)	2.34	2.44
Female (663)	2.87	2.77
eta, beta	.16	.10
Birth Order		
Firstborn (359)	2.87	2.95
Later born (932)	2.51	2.48
eta, beta	.10	.13

NOTE: All eta and beta $P<.001$ (main effect, determined by 4-way ANOVA—classic experimental model) and $R = .646$ and $R^2 = .417$ (MCA, which assumes no interaction). With ANOVA $R = .684$ and $R^2 = .468$ (including interaction). See the text for interaction effects.

birth order. Separate analyses showed that the predictors are related to one another. Using Protestant denomination as the predictor and maternal religiosity (dichotomized) as dependent, the eta for that relationship was .21, with Methodist mothers being least and Baptist and fundamentalist mothers most likely to score high in religiosity.[3] In anticipation of a substantial intercorrelation here, the hierarchical approach was adopted in the ANOVA to ensure that the component sums of squares would add to equal the total sum of squares.

The hierarchical method can be employed if there exists a logical priority among the predictors, whether based in theory or simply on their expected predictive strengths (Overall and Spiegel, 1969:319). Maternal religiosity should be the strongest predictor, but because it is related to denomination, the percentage of variance explained by denomination would decline by entering maternal religiosity first in the hierarchy.[4] It was expected that maternal religiosity would still be the better predictor, even if it followed rather than superseded denomination. Given this and the emphasis placed upon denomination as a variable by Stark and Glock (1968), denomination was first and maternal religiosity second in the ordering. Since I desired to test the efficacy of birth order as a predictor after the effects of the other predictors had been examined, gender was placed third and birth order fourth in the hierarchy. This method of ANOVA estimates each effect adjusted for those before it in the ordering; later effects in the ordering are ignored. The effect of denomination is not adjusted for the other effects, the effect of maternal religiosity is adjusted for denomination, and so on; the interaction effects are adjusted for the main effects.

All main effects are significant at the .001 level. Denomination accounts for 10.5% of the variance, maternal religiosity 28.5%, gender 1.1%, and birth order 1.7%. The main effects account for 41.8% of the variance. The interaction effects account for 5.0% of the variance. The R^2 is .468.[5]

Table 2 presents preadolescent religiosity means by categories of gender, birth order, and denomination. The first hypothesis is that birth-order differences should occur on the part of males but not females, and that an examination of the means by gender and birth order should reveal that females generally display a higher level of religiosity than later-born males. The hypotheses apply to Protestants, not Catholics; the pertinence of this specification can be seen from the second table. Among Protestants, later-born males are less religious than firstborn males. Later-born males and firstborn and later-born females have virtually the same mean on the religiosity index. Examining differences between means by T-tests indicates that for Catholics—specifying gender—there are no significant ($P > .05$) differences by birth order.

The second hypothesis states that birth order differences should be especially pronounced on the part of male youths with mainline, liberal, Protestant identification. Table 2 shows that the interaction between birth order and denomination

[3] The other significant relationship, certainly less important than the one just reported, was between denomination and birth order—eta = .11 and $P < .001$—with Methodists having more firstborn children. This means that their average family size is smaller than for other denominations.

[4] Presumably it would be difficult to disentangle the effects of denomination and parent on the religiosity of youth. This is reflected in Stark's (1972) comment that denomination "indicates the subculture into which a person was born and also something about the kind of religious upbringing he has had." While more religious parents should be overrepresented in the conservative denominations, religiosity varies considerably within denominations.

[5] The following significant ($P < .05$) interactions were present: (1) maternal religiosity and gender (for this effect see Nelsen, 1980; (2) maternal religiosity and denomination (maternal religiosity has a greater impact among youths of the conservative bodies); (3) maternal religiosity, denomination, and birth order; (4) maternal religiosity, denomination, birth order, and gender; (5) gender and birth order (this effect and the next one are central to the hypotheses—we turn to them in the next table); and (6) denomination and birth order. The percent of the variance explained by each was: (1) .3, (2) 1.2, (3) .6, (4) .8, (5) .5, (6) .9. Another .8% was explained by the five other interactions not reaching at least the .05 level of significance. Of interest is that the main effect of birth order explained 1.7% of the variance and the interaction effects including birth order explained an additional 3.1%.

Table 2. Preadolescent Religiosity by Gender, Birth Order, and Denomination

		Religiosity Means by Denomination					
		Protestant				Subtotals	
Gender	Birth Order	Methodist	Episcopal Presbyterian Un. Ch. Christ	Lutheran	Baptist Fundam. Bodies	All Protestant	Catholic
Male	Firstborn	2.48	3.01	2.85	3.28	2.83	2.84
		(43)	(18)	(89)	(23)	(173)	(127)
	Later born	.98	1.96	2.34	3.00	2.16	2.56
		(76)	(40)	(326)	(49)	(491)	(441)
P (T-test)[a]		<.001	<.05	=.01	>.05	<.001	=.07
r		−.47	−.30	−.13	−.08	−.18	−.08
Female	Firstborn	2.23	2.74	2.98	3.43	2.85	2.84
		(44)	(19)	(117)	(23)	(203)	(149)
	Later born	1.71	2.95	2.80	3.92	2.84	3.09
		(70)	(49)	(305)	(79)	(503)	(466)
P (T-test)[a]		<.05	>.05	>.05	=.08	>.05	=.07
r		−.19	.06	−.05	.17	−.002	.07

[a] Two-tailed T-tests for difference between means. Pooled variance estimate used except when an F-test indicated (P<.05) the need for separate variance estimates.

Results of two-way hierarchical ANOVA specifying Protestant grouping (maternal religiosity entered first and birth order second):

Main Effect	Males	Females
Percent of variance explained by prot. denomination	6.7***	11.8***
Percent of variance explained by birth order	3.8***	.1*
Interaction		
Percent explained by denomination by birth order	1.3**	.8*
Total percent of variance explained	11.8	12.7
R	.344	.357

* P>.05 ** P<.05 *** P<.001.

exists in predicting the religiosity of Protestant males but not Protestant females. The differences by birth order for Protestant males are greatest for Methodists, followed by those in the other liberal category, and Lutherans, while the differences by birth order are not significant for Baptist and fundamentalist males.

The hypotheses—(1) that birth order differences should occur for males but not females, with later-born males scoring especially low on the religiosity index; and (2) that an interaction between Protestant denomination and birth order would be present for males, with the most pronounced differences by birth order occurring for those of the mainline denominations—are supported by the data presented in the second table. While the main effect of birth order and the interaction between birth order and denomination lacked significance for Protestant females, the T-test for the difference of birth-order means indicated that Methodist firstborn females were significantly more religious than their later-born counterparts. For this liberal denomination we see birth-order effect for females as well as males.

Earlier, we asked whether birth-order differences in religiosity could be due to a greater percentage of later-born youths in liberal denominations perceiving their mothers as low in religiosity. This is not the case, as can be seen in Table 3 where religiosity means are presented for Protestant males grouped into liberal (Methodist, Episcopal, Presbyterian, and the United Church of Christ) and conservative (Lutheran, Baptist, and fundamentalist) categories. The means are computed for categories of birth order and maternal religiosity, specifying denominational grouping. The later-born males in the liberal grouping score relatively low in religiosity, regardless of their mothers' levels of religiosity. The only anomaly in Table 3 is that liberal, firstborn males who

Table 3. Religiosity of Preadolescent Males by Protestant Grouping, Birth Order, and Maternal Religiosity

		Religiosity Means by Maternal Religiosity		
Protestant Grouping	Birth Order	Low	Moderate	High
Liberal[a]	Firstborn	2.09	3.37	3.24
		(33)	(12)	(17)
	Laterborn	.51	1.81	2.82
		(68)	(14)	(26)
	P (T-test)	<.001	<.01	>.05
	r	−.61	−.55	−.13
Conservative[b]	Firstborn	1.83	3.45	4.35
		(45)	(32)	(27)
	Later born	1.48	2.64	3.94
		(174)	(83)	(99)
	P (T-test)	>.05	<.001	<.05
	r	−.11	−.30	−.13

[a] Methodist, Episcopal, Presbyterian, United Church of Christ.
[b] Lutheran, Baptist, Fundamentalistic Protestant.

Results of two-way hierarchical ANOVA specifying Protestant grouping (maternal religiosity entered first and borth order second):

	Liberal	Conservative
Main Effect	29.7***	40.3***
Percent of variance explained by maternal religiosity	14.0***	1.6*
Percent by birth order		
Interaction		
Maternal religiosity by birth order	2.2**	.2*
Total percent of variance explained	45.9	42.1
R	.678	.649

* P>.05 ** P<.05 *** P<.001.

identify their mothers as high in religiosity score lower in religiosity than would be expected.

The results of the two ANOVA's reported in Table 3 emphasize the differences between liberal and conservative males. Maternal religiosity explains a greater percentage of the variance in the religiosity of conservative than of liberal males. Birth order is a significant predictor for the liberal males (accounting for 14.0% of the variance) but not for the conservative males (it accounts for only 1.6% of the variance). The interaction between maternal religiosity and birth order for the liberal males, while significant, is not substantial (it accounts for 2.2% of the variance) in comparison to the main effects of birth order. The interaction effect between maternal religiosity and birth order lacks significance in the case of conservative males.

There is the special case of only children, the majority of whom were Protestant. The means on the religiosity index were 3.30 (N = 13) and 2.72 (N = 20) for the Protestant males and females, respectively. T-tests for the difference between means indicated no significant (P > .05) differences between male only-children and firstborn males or between female only-children and either firstborn or later-born females. Male only-children were more religious than later-born males (P < .02). Applying this finding to Smith's (1971) logic stated above, we can reject an interpretation of birth-order effect limited to the firstborn being in a superordinate position.

Conclusions

In the review of previous studies, I contrasted the religious climate of two decades ago and the present and observed that the liberal Protestant denominations have fared poorly. Using data from one point in time (1975) I looked at the effects of gender and birth order in producing conformity to parental teaching. Religious socialization should be considered within its context—here consisting of three

nested levels: climate (time), denomination, and family processes (gender and birth order).

In identifying later-born males in the liberal denominations as those youths who especially display lower levels of religiosity, I have focused on a basic discontinuity in the familial transmission of religiosity. From what do these birth-order differences stem? Hoge and Roozen (1979a:58) noted that a decline in religiosity can be associated with a change in values—for example, increased permissiveness in sexual outlook. In this article the interpretation of the effects of birth order and gender is that the firstborn males and females generally tend to be oriented to their parents, whereas later-born males and even some later-born females are not as conforming, perhaps because they have become more peer-oriented. At some points in history the culture of the younger generation can diverge from that of their parents, although, of course, all youths do not participate equally in the change (the level of change varies from being countercultural to representing modifications of the parental culture). At other times, stability in values across generations predominates. When change is occurring, some youths have more access to the producers and transmitters of change, which may include the mass media and peer-group culture of a nonconforming nature.

Events and conditions shape outlook; the 1960s and 1970s saw civil-rights and war protests and the formation of a counterculture, as well as growing cynicism about the institutions of society. The religious disaffection of youth has been charted by Wuthnow (1976), following Mannheim (1952). Observation of birth-order and gender effects especially among youths of the mainline, liberal, Protestant bodies adds contextual understanding of this generational change. Further research should clarify what value shifts associated with the decline in religiosity occurred especially among these youths and whether the patterns associated with this decline in religiosity remain.

We have observed a complex pattern of interaction in religious transmission, involving birth order, gender, and denomination. By considering these and other variables we can go beyond the simple dictum that it is the parents in conjunction with the church that encourage the child to become religious. Stability in religious transmission across generations might, in fact, result from a fairly simple socialization process; discontinuity in religious transmission apparently does not.

REFERENCES

Acock, Alan C. and Vern L. Bengtson
1980 "Socialization and attribution processes: actual versus perceived similarity among parents and youth." Journal of Marriage and the Family 42:501–15.

Argyle, Michael and Benjamin Beit-Hallahmi
1975 The Social Psychology of Religion. London: Routledge and Kegan Paul.

Bibby, Reginald W.
1978 "Why conservative churches *really* are growing: Kelley revisited." Journal for the Scientific Study of Religion 17:129–37.

Carroll, Jackson W., Douglas W. Johnson, and Martin E. Marty
1979 Religion in America: 1950 to the Present. New York: Harper and Row.

DeBord, Larry W.
1969 "Adolescent religious participation: an examination of sib-structure and church attendance." Adolescence 4:557–70.

Gallup Organization
1975 Religion in America. The Gallup Opinion Index, Report 114. Princeton, N.J.: Gallup Organization.

Garai, Josef E. and Amram Scheinfeld
1968 "Sex differences in mental and behavioral traits." Genetic Psychology Monographs 77, Second Half:169–299.

Hoge, Dean R.
1974 Commitment on Campus. Philadelphia: Westminster Press.

Hoge, Dean R. and David A. Roozen
1979a "Research on factors influencing church commitment." Pp. 42–68 in Dean R. Hoge and David A. Roozen (eds.), Understanding Church Growth and Decline: 1950:1978. New York: Pilgrim Press.
1979b "Some sociological conclusions about church trends." Pp. 315–33 in Dean R. Hoge and David A. Roozen (eds.), Understanding Church Growth and Decline: 1950–1978. New York: Pilgrim Press.

Johnson, Benton
1963 "On church and sect." American Sociological Review 28:539–49.

Long, Barbara H. and Edmund H. Henderson
1973 "Children's use of time: some personal and social correlates." Elementary School Journal 73:193–9.

MacDonald, A. P., Jr.
1969a "Manifestations of differential levels of socialization by birth order." Developmental Psychology 1:385–99.

1969b "Birth order and religious affiliation." Developmental Psychology 1:628.

Mannheim, Karl
1952 Essays on the Sociology of Knowledge. New York: Oxford University Press.

Mol, Hans
1971 Religion in Australia. Melbourne: Thomas Nelson.
1976 Identity and the Sacred. Oxford: Basil Blackwell.

Nelsen, Hart M.
1980 "Religious transmission versus religious formation: preadolescent-parent interaction." Sociological Quarterly 21:207–18.

Overall, John E., and Douglas Spiegel
1969 "Concerning least squares analysis of experimental data." Psychological Bulletin 72:311–22.

Putney, Snell and Russell Middleton
1961 "Rebellion, conformity, and parental religious ideologies." Sociometry 24:125–35.

Rogers, Dorothy
1969 Child Psychology. Belmont, California: Brooks/Cole.

Roof, Wade Clark and Dean R. Hoge
1980 "Church involvement in America: social factors affecting membership and participation." Review of Religious Research 21, 4 (Supplement):405–26.

Roozen, David A.
1979 "The efficacy of demographic theories of religious change: Protestant church attendance, 1952–68." Pp. 123–43 in Dean R. Hoge and David A. Roozen (eds.), Understanding Church Growth and Decline: 1950–1978. New York: Pilgrim Press.
1980 "Church dropouts: changing patterns of disengagement and re-entry." Review of Religious Research 21, (Supplement): 427–50.

Sampson, Edward E. and Francena T. Hancock
1967 "An examination of the relationship between ordinal position, personality, and conformity." Journal of Personality and Social Psychology 5:398–407.

Smith, Thomas Ewin
1971 "Birth order, sibship size and social class as antecedents of adolescents' acceptance of parents' authority." Social Forces 50:223–32.

Stark, Rodney
1972 "The economics of piety: religious commitment and social class." Pp. 483–503 in Gerald W. Thielbar and Saul D. Feldman (eds.), Issues in Social Inequality. Boston: Little, Brown.

Stark, Rodney and Charles Y. Glock
1968 American Piety: The Nature of Religious Commitment. Berkeley: University of California Press.

Warren, Jonathan R.
1966 "Birth order and social behavior." Psychological Bulletin 65:38–49.

Wilson, Bryan R.
1959 "An analysis of sect development." American Sociological Review 11:3–15.

Wingrove, C. Ray and Jon P. Alston
1974 "Cohort analysis of church attendance, 1939–1969." Social Forces 53:324–31.

Wuthnow, Robert
1976 "Recent patterns of secularization: a problem of generations?" American Sociological Review 41:850–67.

Yinger, J. Milton
1970 The Scientific Study of Religion. New York: Macmillan.

Zald, Mayer N. and Roberta Ash
1966 "Social movement organizations: growth, decay and change." Social Forces 44:327–41.

SOCIAL CLASS AND ETHNIC SEGREGATION: A RECONSIDERATION OF METHODS AND CONCLUSIONS*

DOUGLAS S. MASSEY
University of Pennsylvania

Previous research on the relationship between social class and ethnic segregation has produced inconsistent results. This paper resolves these inconsistencies by clarifying some theoretical and methodological issues. Ecological theory hypothesizes an inverse relationship between ethnic segregation and socioeconomic status. We demonstrate that this hypothesis cannot be tested using the method of indirect standardization, but is appropriately examined using either direct standardization or cross-sectional correlation. Given these clarifications, previous research generally supports hypotheses derived from ecological theory and lends credence to a social class interpretation of ethnic segregation.

INTRODUCTION

According to ecological theory, socioeconomic advancement by ethnic groups should lead to their progressive integration within American society (Burgess, 1925; McKenzie, 1925; Park, 1926; Duncan and Lieberson, 1959). Based on this perspective, various researchers have posited a "social status" hypothesis specifying a relationship between social class and ethnic segregation (Duncan and Lieberson, 1959; Lieberson, 1963; Darroch and Marston, 1971; Guest and Weed, 1976; Massey, 1979a, 1981a). At the same time, other sociologists have put forth a contrary position emphasizing the persistence of ethnic cohesion over time, in spite of social mobility (Gordon, 1964; Kantrowitz, 1973; Bleda, 1978). Based on this "ethnic status" hypothesis, little relationship between social class and ethnic segregation is predicted.

These two contrasting positions have been tested using three different methodological approaches. The cross-sectional approach correlates degree of ethnic segregation with a group's socioeconomic status or dissimilarity at a point in time. The direct standardization approach measures the degree of interethnic segregation *within* different socioeconomic classes, and then examines changes in the level of segregation across classes. The indirect standardization approach estimates the degree of ethnic segregation that would be expected if two groups were segregated along class, but not ethnic, lines, and then compares this value to the amount of ethnic segregation actually observed.

The empirical evidence generated using these three different methods is inconsistent. Some studies appear to support a social class interpretation of segregation, whereas others seem to confirm the ethnic status hypothesis. While studies have repeatedly reported significant cross-sectional correlations between ethnic segregation and socioeconomic status (Duncan and Lieberson, 1959; Lieberson, 1961, 1963; Darroch and Marston, 1971; Guest and Weed, 1976; Massey, 1979a, 1981a), both Darroch and Marston (1971) and Bleda (1978) found little evidence of a social class effect using the technique of indirect standardization. However, using the method of direct standardization, Massey (1979a) found a pronounced inverse relationship between Hispanic-Anglo segregation and socioeconomic status. In contrast, when Darroch and Marston (1971) employed the same method to study ethnic segregation in Toronto, they found no social class effect; and Kantrowitz (1973) reported no effect among Puerto Ricans in New York.

* Direct all correspondence to Douglas S. Massey, Population Studies Center, University of Pennsylvania, 3718 Locust Walk CR, Philadelphia, PA 19174.

The purpose of this paper is to reconcile these apparent inconsistencies by clarifying some of the theoretical and methodological issues involved in determining the social class basis of ethnic segregation. Given these clarifications, results support the ecological theory of assimilation and its prediction of an inverse relationship between socioeconomic status and ethnic segregation.

UNDERSTANDING THE RELATIONSHIP BETWEEN SOCIAL CLASS AND ETHNIC SEGREGATION: THEORETICAL ISSUES

The fundamental principles of ecological theory imply a relationship between ethnic segregation and social class. According to Park (1926:18), "It is because social relations are so frequently and so inevitably correlated with spatial relations; because physical distances are . . . indexes of social distances, that statistics have any significance whatever for sociology." In practice, this broad perspective has been specified in terms of two related, yet distinct, hypotheses, each drawing its inspiration directly from the work of the Chicago School theorists. In considering the social status hypothesis, it is important to distinguish between these two contrasting points of departure.

One line of theory and research is concerned with the process of ethnic assimilation (cf. Duncan and Lieberson, 1959; Lieberson, 1963; Taeuber and Taeuber, 1964; Jones, 1967; Massey, 1979a). Researchers in this area are specifically interested in patterns of spatial integration between various ethnic minorities and some larger "majority" population (such as native whites of native parentage or non-Hispanic whites). They predict an inverse relationship between social class and ethnic segregation. This hypothesis follows directly from the work of theorists such as Burgess (1925), McKenzie (1925), and Park (1926). According to these theorists, segregation occurs "first, on the basis of language and culture." Ultimately, however,

> other processes of selection inevitably take place which bring about segregation based on vocational interests, upon intelligence, and personal ambition. . . . More and more as the ties of . . . language and of culture are weakened, successful individuals move out and eventually find their places in business and in the professions, *among the older population group which has ceased, to be identified with any language or racial group.* The point is that change of occupation, personal success or failure—changes of economic and social status, in short—tend to be registered in changes of location. [Park, 1926:9; italics added]

Thus, as an ethnic group's socioeconomic status increases, its spatial segregation from the majority should decrease.

Other researchers object to measuring ethnic segregation relative to an ambiguous "majority" population. Instead, they analyze patterns of segregation between specific ethnic groups, typically defined on the basis of foreign birth or parentage, and ignore the larger population of individuals born in this country of native parents (cf. Kantrowitz, 1973; Guest and Weed, 1976). Since assimilation, as defined by the Chicago theorists and as employed here, necessarily involves interaction with some larger majority, these researchers are not theorists of assimilation per se. Rather, they are interested in patterns of spatial interrelation between specific ethnic groups, irrespective of their relationship to the majority. Their purpose is not so much to analyze the effects of socioeconomic factors on spatial assimilation as to account for variation in the degree of segregation between different ethnic groups at a point in time. Following Park (1926), they hypothesize that the degree of physical separation between ethnic groups is a direct function of the social distance between them. Ethnic segregation should therefore be positively associated with socioeconomic differences between groups (Darroch and Marston, 1971; Guest and Weed, 1976; Massey, 1981a).

Correctly interpreted, these two propositions are not contradictory. They simply address different questions. One is concerned with the spatial assimilation of ethnic minorities relative to some majority. The other seeks to explain systematic variation in the degree of interethnic segregation in terms of varying social distances between groups. Although they are

different versions of the same overall "social status" hypothesis, the "assimilation" and "social distance" models are conceptually distinct, and have very different records of support in the research literature. Part of the apparent inconsistency of results concerning the social status hypothesis stems from not distinguishing clearly between these two different points of view.

The social distance hypothesis is appropriately tested by using cross-sectional correlation and regression to examine the relationship between residential and socioeconomic dissimilarity among ethnic groups. In studies using such methods, the hypothesis has found widespread support. In general, social distance is highly correlated with residential dissimilarity between particular ethnic groups (Duncan and Lieberson, 1959; Lieberson, 1963; Darroch and Marston, 1971; Guest and Weed, 1976; Massey, 1981a).

In contrast, the assimilation hypothesis has not received consistent support. Darroch and Marston (1971), Guest and Weed (1976), and Bleda (1978) all indicate that indirect standardization is the appropriate method for testing the assimilation model of ethnic segregation. However, using this method, these studies have found socioeconomic factors to have a negligible impact on ethnic segregation. This result is perplexing given that Massey (1979a) found a very strong inverse relationship between socioeconomic status and Hispanic-Anglo segregation using the method of direct standardization, and that Hispanic-Anglo segregation is highly and negatively correlated with Hispanic SES across urban areas (Grebler et al., 1970; Massey, 1979a). This apparent inconsistency arises from the inappropriate use of the method of indirect standardization, as demonstrated in the following section.

ASSESSING THE SOCIAL CLASS BASIS OF SEGREGATION: METHODOLOGICAL ISSUES

The three techniques mentioned above (cross-sectional correlation, indirect standardization, and direct standardization) all use the index of dissimilarity as their basic measure of segregation. This index measures segregation as the degree of departure from an even residential distribution of minority and majority groups (Duncan and Duncan, 1955a; Taeuber and Taeuber, 1965). Conceptually, it represents the relative proportion of minority members that would have to exchange tracts with majority members to achieve evenness. For example, if blacks were evenly distributed throughout the census tracts of a city, the expected number of blacks in any of j=1 to m tracts would be: $E(B_j)=qP_j$, where q is the city-wide black proportion. The number of blacks required to exchange residences with whites in other tracts to achieve evenness would thus equal B^*-qP^*, where B^* and P^* refer to black and total populations in those tracts where $B_j>E(B_j)$, i.e. those in which blacks are overrepresented. When segregation is at a maximum, B^*-qP^* reduces to BW/P, and the index of dissimilarity represents the ratio of these two values (Jakubs, 1977:282):

$$D = (B^*-qP^*)/(BW/P) \qquad (1)$$

In practice, the index is usually obtained from the following computational formula:

$$D = \tfrac{1}{2} \sum_{j=1}^{m} \left| \frac{B_j}{B} - \frac{W_j}{W} \right| \qquad (2)$$

However, even if there were no discrimination in urban housing markets, an even residential distribution of blacks and whites would not be expected. These groups obviously differ with respect to socioeconomic status, and studies have shown significant residential segregation between social classes, irrespective of race or ethnicity (Duncan and Duncan, 1955b; Uyeki, 1964; Farley, 1977; Simkus, 1978).

If blacks and whites are cross-tabulated by socioeconomic status within census tracts, the effect of social class can be controlled directly by computing the index of dissimilarity within socioeconomic strata:

$$D_i = (B^*_i-q_iP^*_i)/B_iW_i/P_i) \qquad (3)$$

where $q_i = \sum_{j=1}^{m} B_{ij} / \sum_{j=1}^{m} P_{ij}$, and B^*_i and P^*_i are the black and total populations in stratum i within those tracts where $B_{ij}>q_iP_{ij}$. This direct standardization procedure provides

an appropriate test of the assimilation hypothesis that ethnic segregation varies inversely with increasing social class. If spatial assimilation is an inverse function of SES, then we should observe a regular decline in within-class segregation scores as we move up the socioeconomic scale. Such a pattern would be consistent with the strong negative correlation typically observed between SES and ethnic segregation across cities (cf. Massey, 1979a).

Unfortunately, many groups are not cross-tabulated by SES variables at the census tract level, and the direct standardization procedure cannot be applied. However, if the total population is cross-tabulated by social class within tracts, and by race and social class in the city as a whole, the expected number of blacks per tract can be adjusted to indirectly take account of segregation by social class. Specifically, by assuming an even distribution of blacks and whites within social classes, but allowing segregation between classes, the expected number of blacks in tract j becomes $E(B_j) = \sum_{i=1}^{n} q_i P_{ij}$, for i = 1 to n social classes. This expectation yields an index of dissimilarity conditioned by social class (Duncan et al., 1961):

$$D' = (B^* - \sum_{i=1}^{n} q_i P^*_j)/(BW/P) \qquad (4)$$

where B^*_j and P^*_j again refer to black and total populations in those tracts where $B_j > E(B_j)$. This index states the degree of racial segregation not attributable to social class. The degree of racial segregation expected on the basis of social class alone can be computed as a residual:

$$E(D) = D - D' \qquad (5)$$

or it can be calculated directly from the expected distributions of blacks and whites among tracts:

$$E(D) = \tfrac{1}{2} \sum_{j=1}^{m} \left| \frac{\sum_{i=1}^{n} q_i P_{ij}}{B} - \frac{\sum_{i=1}^{n} (1-q_i) P_{ij}}{W} \right| \qquad (6)$$

The method of indirect standardization examines the relationship between D and E(D) to determine the degree to which minority group segregation can be attributed to social class, either by correlating expected and actual numbers of minority members across tracts (Duncan et al., 1961; Taeuber and Taeuber, 1965), or simply by forming the ratio E(D)/D (Darroch and Marston, 1971; Bleda, 1978).

According to Bleda (1978:92), indirect standardization represents a more appropriate means of assessing the role of social class in ethnic segregation than cross-sectional correlation or regression; and, based on unsupportive results using this technique, a social class interpretation of ethnic segregation has been rejected in favor of one based on ethnic status (Darroch and Marston, 1971; Bleda, 1978).

However, indirect standardization does not test the assimilation hypothesis that ethnic segregation declines with increasing socioeconomic status. It is based on a two-stage model of ethnic segregation where (1) groups differ in terms of socioeconomic status, and (2) socioeconomic classes are segregated from one another. While it does measure the relative contribution of social class to ethnic segregation, it does not address the issue of whether ethnic segregation is inversely related to socioeconomic status. As demonstrated below, it is entirely plausible for social class variables to account for a small proportion of ethnic segregation, but for ethnic segregation simultaneously to show a pronounced decline with increasing social status. Thus, given a small social class effect using indirect standardization, one cannot conclude that social class has no effect on segregation, and one cannot reject the assimilation hypothesis in favor of one based solely on ethnic solidarity.

Indirect standardization attempts to explain ethnic segregation in terms of certain "structural" characteristics of the city under consideration: the socioeconomic distributions of minority and majority groups, and the degree of social class segregation. Social class will tend to account for a small share of residential segregation either if the two groups have similar socioeconomic characteristics or if there is relatively little segregation by social status, irrespective of the actual degree of residential segregation within so-

Table 1. Distribution of Blacks among Census Tracts and Social Classes in a Hypothetical City

Social classes		Tracts 1	2	3	4	Total	
1	White	45	45	5	5	100	
	Black	20	20	180	180	400	$q_1 = .8$
	Total	65	65	185	185	500	
2	White	90	90	10	10	200	
	Black	15	15	135	135	300	$q_2 = .6$
	Total	105	105	145	145	500	
3	White	135	135	15	15	300	
	Black	10	10	90	90	200	$q_3 = .4$
	Total	145	145	105	105	500	
4	White	180	180	20	20	400	
	Black	5	5	45	45	100	$q_4 = .2$
	Total	185	185	65	65	500	
Total	White	450	450	50	50	1000	
	Black	50	50	450	450	1000	$q = .5$
	Total	500	500	500	500	2000	

$$D = \frac{900 - (.5)(1000)}{(1000)^2/2000} = .80 \qquad \text{From (1).}$$

$$E(D) = \frac{1}{2}\left[\left|\frac{290-210}{1000}\right| + \left|\frac{290-210}{1000}\right| + \left|\frac{210-290}{1000}\right| + \left|\frac{210-290}{1000}\right|\right] = .16 \qquad \text{From (6)}$$

$$E(D)/D = .20$$

$$D_1 = \frac{360 - (.8)(370)}{(400)(100)/500} = .80$$

$$D_2 = \frac{270 - (.6)(290)}{(300)(200)/500} = .80$$

$$D_3 = \frac{180 - (.4)(210)}{(200)(300)/500} = .80 \qquad \text{From (1)}$$

$$D_4 = \frac{90 - (.2)(130)}{(100)(400)/500} = .80$$

cioeconomic classes. Thus, given the fact that the degree of residential segregation between socioeconomic classes is not particularly impressive (cf. Duncan and Duncan, 1955a; Uyeki, 1964; Farley, 1977; Simkus, 1978), it is not surprising that social class has been found to account for a relatively small share of ethnic segregation.

The limitations of indirect standardization as a method for studying the effect of social class on segregation are clearly revealed by a hypothetical example. Tables 1 and 2 present idealized "black" and "Hispanic" populations residing in two hypothetical cities. These cities each contain the same number of residents, have the same minority proportion, the same number and sizes of census tracts, and the same distribution of minority and majority by socioeconomic status. Segregation is equally high in both cities, with indices of dissimilarity equal to .80. Moreover, controlling for the effect of social class through indirect standardization leads to essentially the same conclusion for both minority groups. The ratio of expected to actual segregation is .20 for blacks and .28 for Hispanics. Such ratios have led other researchers to reject the assimilation hypothesis with respect to ethnic or racial segregation (cf. Taeuber and Taeuber, 1965; Darroch and Marston, 1971; Bleda, 1978).

Table 2. Distribution of Hispanics among Census Tracts and Social Classes in a Hypothetical City

Social classes		Tracts 1	2	3	4	Total	
1	White	45	45	5	5	100	
	Hispanic	0	0	200	200	400	$q_1 = .8$
	Total	45	45	205	205	500	
2	White	90	90	10	10	200	
	Hispanic	0	0	150	150	300	$q_2 = .6$
	Total	90	90	160	160	500	
3	White	135	135	15	15	300	
	Hispanic	25	25	75	75	200	$q_3 = .4$
	Total	160	160	90	90	500	
4	White	180	180	20	20	400	
	Hispanic	25	25	25	25	100	$q_4 = .2$
	Total	205	205	45	45	500	
Total	White	450	450	50	50	1000	
	Hispanic	50	50	450	450	1000	$q = .5$
	Total	500	500	500	500	2000	

$$D = \frac{900 - (.5)(1000)}{(1000)^2/2000} = .80 \qquad \text{From (1)}$$

$$E(D) = \frac{1}{2}\left[\left|\frac{305 - 195}{1000}\right| + \left|\frac{305 - 195}{1000}\right| + \left|\frac{195 - 305}{1000}\right| + \left|\frac{195 - 305}{1000}\right|\right] = .22 \qquad \text{From (6)}$$

$$E(D)/D = .28$$

$$D_1 = \frac{400 - (.8)(410)}{(400)(100)/500} = .90$$

$$D_2 = \frac{300 - (.6)(320)}{(300)(200)/500} = .90$$

$$D_3 = \frac{150 - (.4)(180)}{(200)(300)/500} = .65 \qquad \text{From (1)}$$

$$D_4 = \frac{50 - (.2)(90)}{(100)(400)/500} = .40$$

However, when direct standardization is performed, a very different picture emerges. That is, when social class is controlled directly by computing the index of dissimilarity within classes, Hispanics show a marked decrease in segregation from whites with greater socioeconomic status, while blacks do not. Hispanic-white segregation declines from a high of .90 in the lowest socioeconomic class to .40 in the highest, while black-white segregation is invariant at .80 across all social classes. In other words, patterns of Hispanic-white segregation support the assimilation hypothesis, even though indirect standardization shows social class to account for a relatively small share of ethnic segregation.

This seeming contradiction results because indirect standardization is based on expectations derived solely from the marginal distributions of Tables 1 and 2. It does not take into account the distribution of minority and majority groups within cells of the table formed by cross-classifying tracts, social class, and ethnicity, and therefore it reveals nothing concerning patterns of ethnic segregation within social classes. Given a city with m tracts, n social classes, and two groups, there are (m−1)(n−1) degrees of freedom in the resulting cells. Thus even with tract sizes and socioeconomic distributions fixed—and hence the overall level of segregation and the amount due to social class determined—there are still a wide

variety of possible patterns of within-class ethnic segregation.

The former hypothetical examples were presented for heuristic purposes; but the points illustrated have real relevance with respect to the kind of empirical results one obtains in practice, and consequently to the theoretical inferences one draws. Table 3 replicates the exercises of Tables 1 and 2 using actual data from six southwestern SMSAs. In this example, standardization is on the basis of education.

Expected values of the index of dissimilarity are quite small for both Hispanic-white and black-white segregation. That is, the relative share of each group's segregation that can be attributed to education is quite small. The ratios of E(D) to D are relatively larger for Hispanics because they are much less segregated from whites than are blacks. Nonetheless, the ratios for Hispanics cluster closely around a value of .20, and based on similar findings, both Darroch and Marston (1971) and Bleda (1978) have previously rejected a socioeconomic interpretation of ethnic segregation. Were the analysis to end at this point, we might conclude that social class has a negligible effect on Hispanic as well as black segregation.

Table 3. Observed Dissimilarity Indices and Values Expected Assuming Social Class Segregation, Six Southwestern SMSAs, 1970

	Hispanic-White			Black-White		
	D	E(D)	E(D)/D	D	E(D)	E(D)/D
Los Angeles	46.6	10.1	.217	90.0	6.8	.076
San Francisco	35.2	7.2	.205	76.9	7.5	.098
San Diego	32.3	7.9	.244	83.4	6.7	.080
Denver	47.2	8.6	.182	87.2	3.9	.045
Sacramento	34.8	6.6	.190	66.5	5.4	.081
Tucson	53.2	11.7	.185	67.5	7.4	.110
Average	41.6	8.7	.204	78.6	6.3	.082

However, the results of direct standardization (depicted in Figure 1) suggest a far different conclusion. In each of the SMSAs, Hispanic-white segregation displays a striking decrease with increasing education, while in most cases black-white segregation tends to be invariant across educational strata. Rather than questioning the assimilation hypothesis derived from ecological theory, patterns of Hispanic-white segregation provide a rather eloquent confirmation of it.

Thus indirect standardization represents an inappropriate test of the assimilation hypothesis. While it can be validly used to measure the degree of ethnic segregation attributable to social class, it contributes nothing towards understanding how socioeconomic variables influence ethnic segregation, the issue of true interest in studying processes of assimilation among ethnic groups.

This distinction has probably remained obscure because conclusions based on indirect standardization were coincidentally correct for the first group to which it was applied (i.e. blacks). Thus the Taeubers' (1965:95) assertion, based on indirect standardization, that "improving the economic status of Negroes is unlikely by itself to alter prevailing patterns of racial residential segregation" was correct, as Farley (1977) proved using direct standardization. However, in the case of Hispanics, the results of indirect standardization would be very misleading, and such a conclusion would be incorrect. While indirect standardization produces results very similar to those of blacks, one finds a very powerful social class effect using direct standardization. Among Hispanics, improvements in socioeconomic status would very likely lead to lower levels of residential segregation.

ETHNIC SEGREGATION AND SOCIAL CLASS RECONSIDERED

We have demonstrated that finding a small social class effect using indirect standardization is consistent with the assimilation hypothesis. On the other hand, Kantrowitz (1973) obtained unsupportive results using direct standardization, thus undermining the assimilation hypothesis. However, my previous work (Massey, 1979a, 1979b, 1981a) has already shown that New York's Puerto Ricans (the group examined by Kantrowitz) represent a clear exception to overall patterns of Hispanic segregation, which generally show a strong inverse relationship between

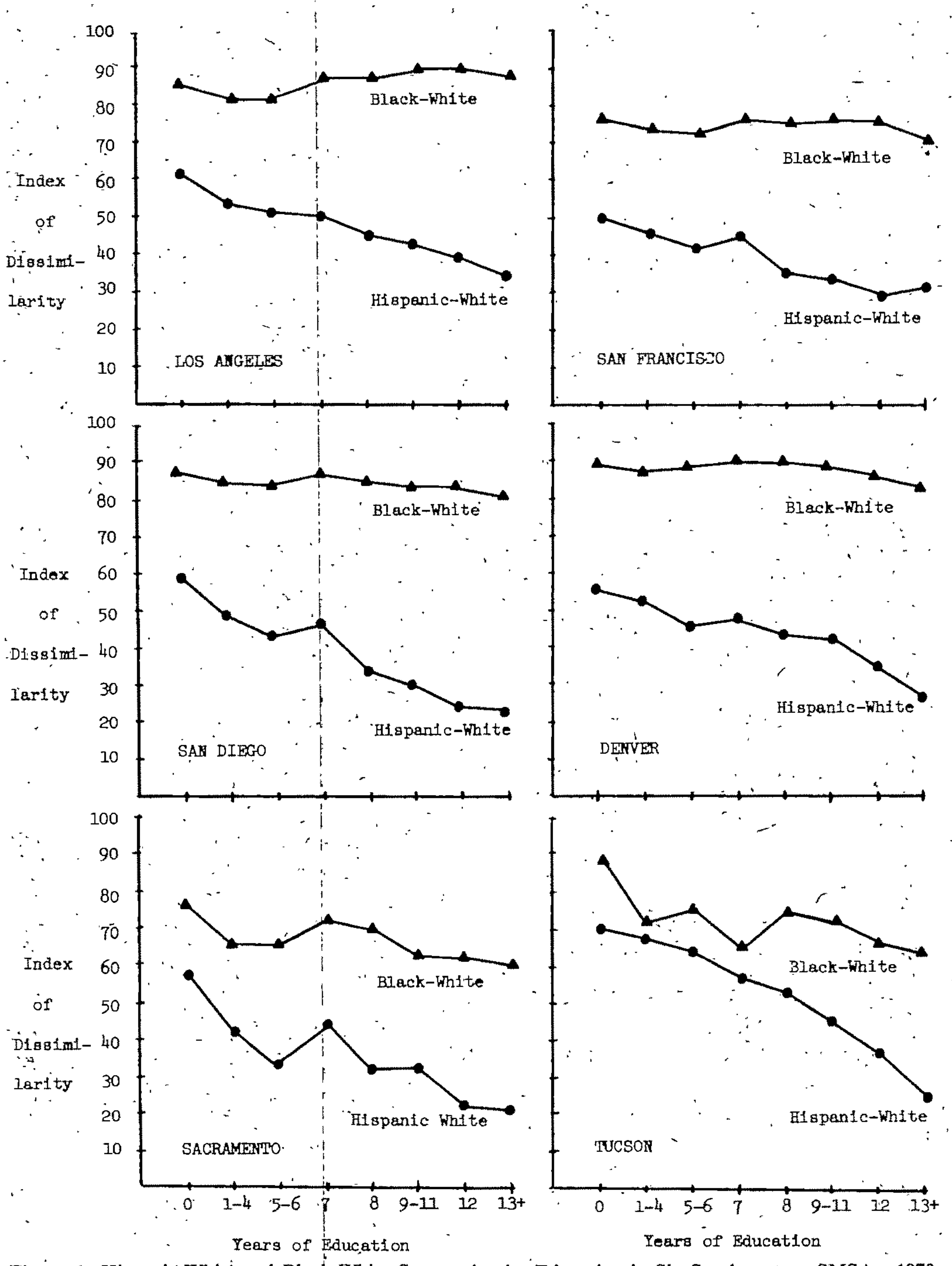


Figure 1. Hispanic-White and Black-White Segregation by Education in Six Southwestern SMSAs, 1970

segregation and social class. Indeed, New York's Puerto Ricans score lower than Hispanic groups in other cities on a variety of different measures of assimilation (1981b), and thus represent more of an anomaly than a definitive contradiction of the assimilation hypothesis. Indeed, with respect to processes of assimilation, Puerto Ricans behave more like a racial than an ethnic group, possibly because of the large black admixture in this population (Seda Bonilla, 1972; Longres, 1974).

Darroch and Marston (1971) also obtained unsupportive results using the

method of direct standardization in Toronto. No consistent differences were found in the level of ethnic segregation between white-collar and blue-collar or low- and high-income groups. A possible explanation for these findings may relate to the use of mother tongue and birth place to define ethnic groups. The use of foreign birth to define ethnicity excludes from consideration all second and third generation immigrants. Given that processes of assimilation require time to operate, a single generation may be insufficient to observe the hypothesized effects of social mobility. Definition of ethnicity based on mother tongue could have the same effect, if reporting a mother tongue other than English were highly related to foreign birth.

In general, at this point the evidence is insufficient to reject the assimilation hypothesis. Cross-sectional correlations between level of ethnic segregation and socioeconomic status have consistently been significant and in the expected direction. At the same time, detailed analyses of patterns of Hispanic segregation in the United States reveal an almost monotonic decline in within-class segregation as status increases. Given the problems associated with indirect standardization, the anomalous nature of results from New York and Toronto, and recognizing that there are no good a priori grounds for assuming Hispanics are different from other ethnic groups, the weight of accumulated evidence supports the assimilation hypothesis and its postulated effects on ethnic segregation.

SUMMARY AND CONCLUSION

This paper has attempted to clarify salient theoretical and methodological issues concerning the relationship between social class and ethnic segregation. Theoretically, a distinction was introduced between two variants of the "social status" hypothesis: the assimilation hypothesis, which predicts greater spatial integration between minority and majority groups with increasing social class, and the social distance hypothesis, which posits a direct relationship between physical and social distance among ethnic groups. Methodologically, this paper has demonstrated how results obtained using the method of indirect standardization cannot be used to refute the assimilation hypothesis. Given these clarifications, research generally supports the assimilation hypothesis in its prediction of a negative relationship between social class and ethnic segregation.

However, accepting the assimilation hypothesis does not necessarily imply rejection of the ethnic status hypothesis. Indeed, ethnic cohesion of the mechanical type is clearly an important factor in the persistence of ethnic segregation over time (Suttles, 1968), and helps to explain why ethnic segregation exists even at the highest socioeconomic levels (Massey, 1979a). Nonetheless, it seems clear that increasing social status has an important negative effect on ethnic segregation, which if projected into the future implies the progressive spatial assimilation of ethnic populations over time.

With respect to future research strategies, this paper suggests direct standardization to be the preferred method for assessing the effect of social class on ethnic segregation. Until recently, this preference would have been of little practical utility, since very few ethnic groups were cross-tabulated by socioeconomic variables at the tract level. For example, the 1970 Census provided such information only for blacks, whites, and Spanish Americans. However, plans announced for the 1980 Census indicate that in the future, this situation will be much improved (U.S. Bureau of the Census, 1979). According to current plans, tract-level cross-tabulations by SES will be available for six ethnic groups (English, French, German, Italian, Polish, and Irish) and for four Hispanic origin groups (Mexicans, Cubans, Puerto Ricans, and "other Spanish"). Moreover, tract-level tabulations by SES will also be provided for four additional ethnic groups whose identity will change from SMSA to SMSA, depending on which ethnic groups predominate in a particular area. Thus a more definitive resolution of the issues discussed here will be possible in the near future.

At this point, however, cross-sectional

analyses of the relationship between ethnic segregation and socioeconomic status represent a valid research strategy. To the extent that measures of a group's overall socioeconomic status and indices of its residential dissimilarity both reflect the underlying socioeconomic distribution, a relationship between SES and ethnic segregation is predicted. As a test of the assimilation hypothesis, the cross-sectional method is clearly more appropriate than the method of indirect standardization heretofore employed.

REFERENCES

Bleda, Sharon Estee
1978 "Intergenerational differences in patterns and bases of ethnic residential dissimilarity." Ethnicity 5:91–107.

Burgess, Ernest W.
1925 "The growth of the city: and introduction to a research project." Pp. 47–62 in R. E. Park and E. W. Burgess (eds.), The City. Chicago: University of Chicago Press.

Darroch, A. Gordon and Wilfred G. Marston
1971 "The social class basis of ethnic residential segregation: the Canadian case." American Journal of Sociology 77:491–510.

Duncan, Otis Dudley, Ray P. Cuzzort, and Beverly Duncan
1961 Statistical Geography: Problems of Analyzing Areal Data. New York: Glencoe.

Duncan, Otis Dudley and Beverly Duncan
1955a "A methodological analysis of segregation indexes." American Sociological Review 20:210–7.
1955b "Residential distribution and occupational stratification." American Journal of Sociology 60:493–503.

Duncan, Otis Dudley and Stanley Lieberson
1959 "Ethnic segregation and assimilation." American Journal of Sociology 64:364–74.

Farley, Reynolds
1977 "Residential segregation in urbanized areas of the United States in 1970: an analysis of social class and racial differences." Demography 14:497–518.

Gordon, Milton M.
1964 Assimilation in American Life. New York: Oxford University Press.

Grebler, Leo, Joan W. Moore, and Ralph C. Guzman
1970 The Mexican American People: The Nation's Second Largest Minority. New York: Free Press.

Guest, Avery M. and James A. Weed
1976 "Ethnic residential segregation: patterns of change." American Journal of Sociology 81:1088–111.

Jakubs, John F.
1977 "Residential segregation: the Taueber Index reconsidered." Journal of Regional Science 17:281–303.

Jones, F. Lancaster
1967 "Ethnic concentration and assimilation: an Australian case study." Social Forces 45:412–23.

Kantrowitz, Nathan
1973 Ethnic and Racial Segregation in the New York Metropolis. New York: Praeger.

Lieberson, Stanley
1961 "The impact of residential segregation on ethnic assimilation." Social Forces 40:52–7.
1963 Ethnic Patterns in American Cities. New York: Free Press

Longres, John F., Jr.
1974 "Racism and its effects on Puerto Rican continentals." Social Casework 55:67–75.

McKenzie, Roderick D.
1925 "The ecological approach to the study of the human community." Pp. 63–79 in R. E. Park and E. W. Burgess (eds.), The City. Chicago: University of Chicago Press.

Massey, Douglas S.
1979a "Effects of socioeconomic factors on the residential segregation of blacks and Spanish Americans in U.S. urbanized areas." American Sociological Review 44:1015–22.
1979b "Residential segregation of Spanish Americans in United States Urbanized Areas." Demography 16:653–64.
1981a "Hispanic residential segregation: a comparison of Mexicans, Cubans, and Puerto Ricans." Sociology and Social Research 65:311–22.
1981b "Dimensions of the new immigration to the United States and the prospects for assimilation." Annual Review of Sociology 7:57–85.

Park, Robert E.
1926 "The urban community as a spatial pattern and a moral order." Pp. 3–18 in E. W. Burgess (ed.), The Urban Community. Chicago: University of Chicago Press.

Seda Bonilla, Eduardo
1972 "El problema de la identidad de los Niuyoricans." Revista de Ciencias Sociales 16:107–20.

Simkus, Albert A.
1978 "Residential segregation by occupation and race in ten urbanized areas, 1950–1970." American Sociological Review 43:81–93.

Suttles, Gerald D.
1968 The Social Order of the Slum. Chicago: University of Chicago Press.

Taeuber, Karl E. and Alma Taeuber
1964 "The Negro as an immigrant group: recent trends in racial and ethnic segregation in Chicago." American Journal of Sociology 69:374–82.
1965 Negroes in Cities. Chicago: Aldine.

U.S. Bureau of the Census
1979 1980 Census of Population and Housing: Tentative Publication and Computer Tape Program. Washington, D.C.: U.S. Government Printing Office.

Uyeki, E. S.
1964 "Residential distribution and stratification, 1960–1960." American Journal of Sociology 59:491–8.

THE DIMENSION(S) OF DEMOCRACY: FURTHER ISSUES IN THE MEASUREMENT AND EFFECTS OF POLITICAL DEMOCRACY*

KENNETH A. BOLLEN
Warren, Michigan

BURKE D. GRANDJEAN
University of Virginia

We test whether Bollen's (1980) index of political democracy is bidimensional or unidimensional. Theoretical considerations suggest that political liberty and popular sovereignty are distinct dimensions, but our confirmatory factor analysis reveals that they are not empirically separable. Regression analysis also supports this conclusion, and at the same time provides evidence on the controversial relationship between democracy and inequality.

The problem of conceptualization and measurement has been called "the major obstacle to integrating theory and research" in sociology (Blalock, 1979:881). The crux of the concept/measurement problem is often a question of dimensionality: is a given theoretical construct—alienation, bureaucracy, or political democracy, for instance—unidimensional, bidimensional, or multidimensional?[1]

The question of dimension is crucial to theory because different dimensions of a construct might be expected to have different causes and consequences. In research, measurement operations that do not accurately reflect the dimensions of a construct yield ambiguous results. If a construct is multidimensional, a unitary measure may tap only one dimension, or may confound a number of partially countervailing dimensions. Conversely, an attempt to use separate measures for a unidimensional phenomenon will result in a futile battle with multicollinearity (cf. Blalock, 1963). The dimensionality question is crucial to the integration of theory and research, because it draws attention to the implicit theoretical assertions in any operational definition, and therefore emphasizes the need to incorporate explicit measurement models in causal models.

This paper reconsiders a measure of political democracy recently proposed in *ASR* by Bollen (1979, 1980). Although he identifies two components of political democracy theoretically, Bollen's empirical measure combines these dimensions. Here, we provide (1) a direct, "internal" test of the dimensionality of the measure, based on a confirmatory factor analysis (Jöreskog and Sörbom, 1978) of the relationships among its constituent indicators; and (2) an "external" test, using standard regression techniques to examine the relationship of income inequality to political democracy and its presumed subdimensions. Our results support the unidimensionality of political democracy, and thus confirm and extend Bollen's analyses. Our findings also shed new light on the controversial democracy/inequality relationship. Finally, the distinction between the theoretical and the empirical dimensionality of a construct is clearly demonstrated, and we suggest that the distinction might be usefully applied to other sociological constructs.

POLITICAL DEMOCRACY: A UNI- OR BI-DIMENSIONAL MEASURE?

In its most general sense, "democracy" is almost certainly multidimensional. For

---

* Direct correspondence to Kenneth A. Bollen, 29960 MacKenzie Circle, Warren MI 48092.

Computer time for the anlaysis was provided in part by the Interuniversity Consortium for Political and Social Research at the University of Michigan. We thank the members of the 1980 I.C.P.S.R. seminar on Models with Unmeasured Variables for their suggestions. This effort is collaborative; the authors' names are listed alphabetically.

[1] On alienation, see, e.g., Seeman (1959) and Kohn (1976). On bureaucracy, see Frisbie (1975), Gold (1964), Hall (1963), and Pugh et al. (1968). A useful review of both of these literatures is provided by Bonjean and Grimes (1970). On democracy, see Bollen (1979, 1980) and the sources cited therein, especially Cutright (1963) and Jackman (1975).

example, Hewitt (1977) uses the term to encompass both "political democracy" and "social democracy"—i.e., socialism. Yet as he stresses (1977:451, 451 n), "no assumption is made that these . . . are related. . . . The correlation between democratic experience and socialist strength is only 0.30." Democracy has also been taken by some, especially in early research on the topic (e.g., Cutright, 1963; Lerner, 1958; see also Stack, 1979), to imply stability in government, high rates of voter participation, or both. The criticisms of these usages presented by Bollen (1980) and others can be interpreted as arguments that political democracy, stability, and participation do not comprise a single dimension of variation.

Even with respect to political democracy more narrowly defined, two analytically distinct dimensions can be identified. Bollen (1980:372) defines political democracy as "the extent to which the political power of the elite is minimized and that of the nonelite is maximized." He suggests that its two main components are *popular sovereignty* and *political liberties*. The first of these concerns the extent to which the elites of a country are accountable to the nonelites. Popular sovereignty implies free elections, as wide a franchise as possible, equal weighting of votes, and fair electoral processes. The second dimension, political liberties, concerns institutions through which nonelites can influence the decisions of elites, including free speech, a free press, and freedom of opposition.

The theoretical reasons for regarding these two dimensions as distinct seem sound. In fact, Dahl (1956), Downs (1957), and Lipset (1963) might reserve the term political democracy for the popular-sovereignty component alone, since their discussions of the concept emphasize electoral processes. Nations do seem to vary somewhat independently along the two dimensions. For example, Britain and the United States appear nearly equal with respect to the democracy of their electoral processes. However, the British press is often regarded as less free (or more responsible) in reporting on sensitive government activities (see Wilensky, 1967). There is also some evidence that the two dimensions are even distinct in the collective conscience—that attitudes about issues of popular sovereignty need not be strongly related to attitudes about civil liberties (Weil, 1980).

Nevertheless, Bollen's (1979, 1980) empirical analyses ignore this theoretical distinction. He combines three indicators of popular sovereignty and three of political liberties in a single index. Using confirmatory factor analysis Bollen (1980) shows that a single-factor solution with correlated measurement errors among some of the six indicators fits the observed data well. But he does not consider a two-factor solution at all, and thus neglects the possibility that a bidimensional model might provide a better fit and perhaps eliminate the need for correlated errors.

If these indicators are actually bidimensional, then combining them in a single index could be seriously misleading. For instance, the effect of political democracy on income inequality is a matter of much debate (see, e.g., Cutright, 1967; Hewitt, 1977, 1979; Jackman, 1975, 1980; Lenski, 1966; Rubinson and Quinlan, 1977; Stack, 1978, 1979; Wilensky, 1975). Suppose that through fair electoral processes a population is able to introduce policies which reduce income inequality, but that political liberties alone have little effect on the income distribution. If indicators of both popular sovereignty and political liberties are combined in an index, these relationships may be blurred, possibly leading to null results. Our purpose is to determine if two dimensions can be empirically distinguished that fit Bollen's (1980) six indicators of political democracy better than a unidimensional solution does.

DATA AND METHODS

Bollen's three indicators of political liberty are (X_1) press freedom, (X_2) freedom of group opposition, and (X_3) extent of government negative sanctions. The three popular sovereignty indicators are (X_4) fairness of elections, (X_5) method of executive selection, and (X_6) method of legislative selection (and legislative effectiveness). We follow Bollen's (1979, 1980) definitions and sources of these measures.

As does Bollen, we analyze the 1960 covariance matrix of these six indicators for 113 countries (Bollen, 1980:fn. 27).

First, we estimate a model with two factors—one underlying the political-liberty indicators, the second for the popular-sovereignty measures. To establish the units of the latent factors, they are arbitrarily constrained to have a variance of 1.0, but the correlation between the two factors is unconstrained. A second model is estimated identical to the first except that the correlation between the two factors is constrained to equal 1.0. That is, the two factors are assumed to be perfectly correlated—to be in effect a single factor. (See Jöreskog (1978) for another example of testing whether two scores are perfectly correlated.)

The χ^2 values from the two models may be compared to assess the difference in fit. The second model, with the correlation between the two factors constrained to 1.0, represents a more restricted model than the first, where the correlation between the factors is unconstrained. When two models are nested, the difference between the χ^2 values is itself distributed as χ^2. If a "significant" difference in the χ^2 values for these models is found, we will conclude that the constrained model can be rejected and the two factor solution should be used. On the other hand, a finding of no significant difference between these two models would suggest that a perfect correlation between the factors is an acceptable constraint, and hence that a single factor may underlie all six indicators. Of course, we also want a model that has a good overall fit in addition to having a better fit than some alternative one. If the χ^2 for a model is significant, the fit of the model is poor,[2] and other restrictions on the model should be removed to improve the fit.

We also present a regression-based test of the dimensionality of political democracy. For ease of exposition we defer discussing the methods of that analysis until later in the paper.

AN INTERNAL TEST

We estimate a model with two factors, one for the three political liberty indicators, the other for the three popular sovereignty measures. Initially the errors of measurement are assumed to be uncorrelated, as shown in Figure 1a. The models in Figures 1b and 1c permit correlated error. Following Jöreskog's (1978) notational conventions,[3] a circle indicates a true unmeasured variable or latent factor (ξ_1), a box encloses a fallible indicator (X_1) and measurement error is represented by δ_1.

The models on the left-hand side of the figure are bidimensional. They have the variance of ξ_1 and ξ_2 set to 1.0, but the correlation $\rho_{\xi_1\xi_2}$ unconstrained. The models on the right-hand side are identical except that $\rho_{\xi_1\xi_2}$ is constrained to equal 1.0. The right-hand models are equivalent to assuming a single factor. That is, the χ^2 values are identical to the χ^2 values obtained by Bollen (1980) assuming a unidimensional model.[4]

The results of estimating the six models in the figure are reported in Table 1. The lambda values (or "factor loadings") are similar for each pair of models. For example, the effects of ξ_1 on X_1 for the models in Figure 1a are estimated to be 8.75 and 8.87 for the bidimensional and unidimensional solutions. However, the comparison of greatest interest is the overall fit of each model as measured by χ^2. The difference in χ^2 values in Figure 1a is 6.40, which would be statistically significant in a random sample of this size ($p < 0.05$). This suggests that the two factor solution

[2] An exception to this is if the sample is extremely large. Since χ^2 increases with sample size, statistically significant χ^2s may be found that are of little practical significance. Since we have only 113 cases this is not a problem in our analysis. Note that throughout the paper we treat a test of significance as a useful decision-rule even though our "sample" of nations is simply all countries for which the necessary data are available. This is true for our χ^2 tests as well as the t-statistics calculated by comparing regression coefficients to their estimated standard errors.

[3] Bollen (1980) uses η, Y, and ϵ where we use ξ, X, and δ. For simple factor analytic models the two sets of symbols are interchangeable. We use the latter because constraining the correlations between η variables is notationally more cumbersome.

[4] See Bollen's (1980) Table A1, model A1, A2, and A3. The slight discrepancy for his model A2 is due to a typographical error. The correct χ^2 is our 4.24.

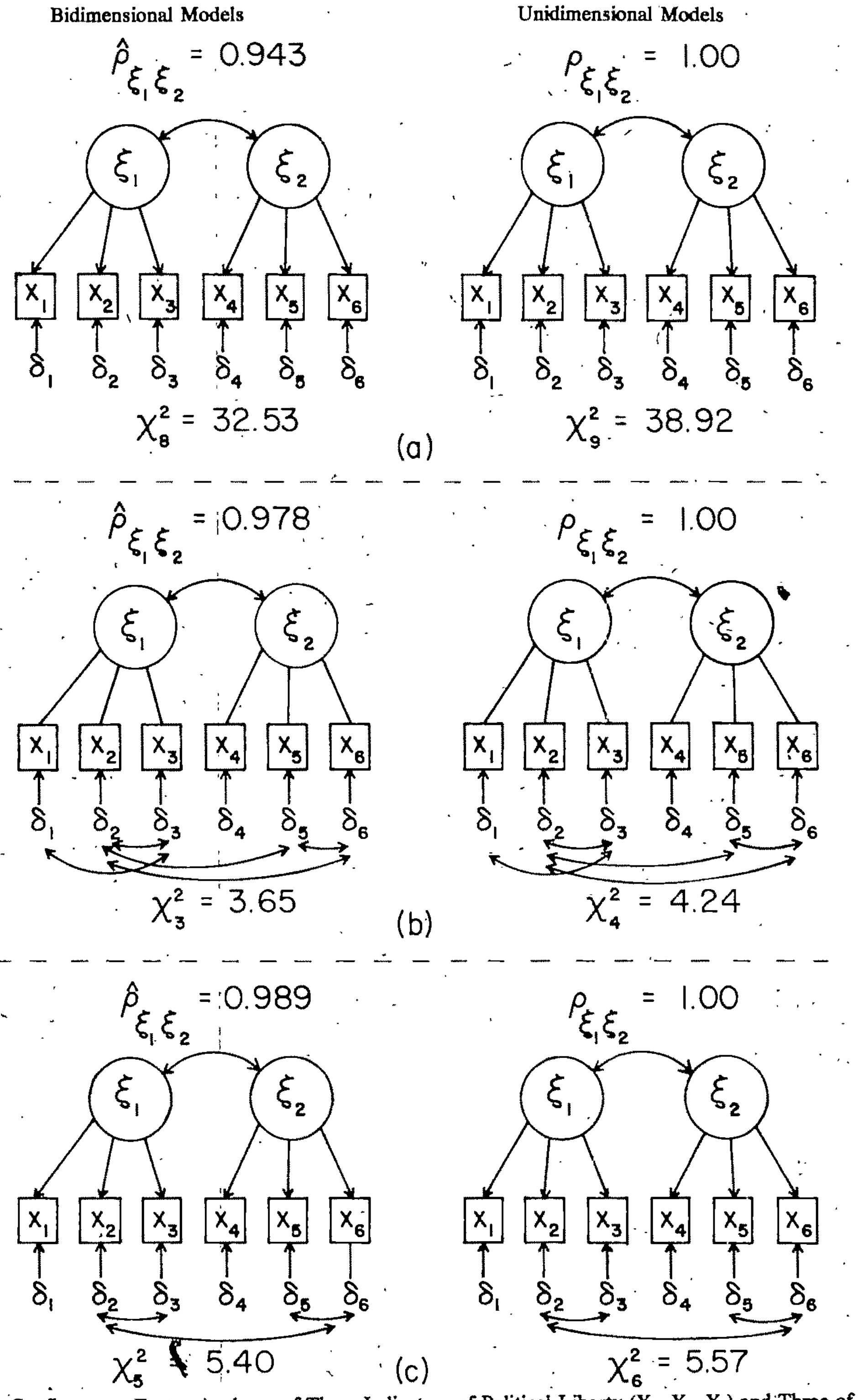


Figure 1. Confirmatory Factor Analyses of Three Indicators of Political Liberty (X_1, X_2, X_3) and Three of Popular Sovereignty (X_4, X_5, X_6), 1960 (N=113)

NOTE: To facilitate statistical comparisons of the models, the unidimensional solution is represented in an equivalent form as two perfectly correlated factors ($\rho_{\xi_1\xi_2} = 1.00$ by constraint). See text for a discussion.

Table 1. Estimates of Confirmatory Factor Analyses for Bidimensional and Unidimensional Models Represented in Figure 1

	Estimates					
	(a)		(b)		(c)	
Parameter	Bi-dimensional	Uni-dimensional	Bi-dimensional	Uni-dimensional	Bi-dimensional	Uni-dimensional
λ_1	8.75	8.87	9.41	9.31	9.24	9.22
λ_2	12.29	12.08	11.13	11.17	11.20	11.19
λ_3	5.27	5.18	5.08	5.03	4.95	4.95
λ_4	8.18	8.14	8.52	8.45	8.57	8.52
λ_5	10.11	9.64	9.60	9.50	9.61	9.55
λ_6	10.24	9.88	9.59	9.51	9.61	9.56
$V(\delta_1)$	23.93	21.90	11.98	13.90	15.06	15.54
$V(\delta_2)$	31.05	36.22	56.53	56.12	54.78	55.31
$V(\delta_3)$	5.33	6.22	7.30	7.86	8.60	8.62
$V(\delta_4)$	39.51	40.21	33.89	35.01	33.06	33.88
$V(\delta_5)$	95.46	104.69	105.56	107.39	105.36	106.57
$V(\delta_6)$	22.75	30.09	35.83	37.29	35.95	36.82
$C(\delta_1\delta_3)$	0.00†	0.00†	$-2.39^{n.s.}$	$-1.55^{n.s.}$	0.00†	0.00†
$C(\delta_2\delta_3)$	0.00†	0.00†	9.42	10.07	10.46	10.65
$C(\delta_2\delta_5)$	0.00†	0.00†	$-1.43^{n.s.}$	$-3.15^{n.s.}$	0.00†	0.00†
$C(\delta_2\delta_6)$	0.00†	0.00†	15.02	13.39	12.97	12.68
$C(\delta_5\delta_6)$	0.00†	0.00†	16.19	17.85	17.28	18.38
$\rho_{\xi_1\xi_2}$	0.94	1.00†	0.98	1.00†	0.99	1.00†
χ^2	32.53	38.92	3.65	4.24	5.40	5.57
Degrees of freedom	8	9	3	4	5	6

NOTE: All parameter estimates are significant at the .05 level except those marked † or n.s. (see footnote 2 on significance tests).

† Constrained parameter.

n.s. Not significant at .05 level.

has a better fit than the single factor model, even though the correlation between the factors is very high (.94). However, each model has a highly significant χ^2 ($p < .0001$). This indicates that neither model does an adequate job of fitting the data. Thus, before we prematurely conclude that a two factor model is preferable to a single factor, it is first necessary to find a model with a good fit.

Bollen (1980) suggests that one or more of the errors of measurement for the indicators may be correlated, contrary to our initial assumption in Figure 1a. In particular the errors of measurement of the variables coming from the same data source (i.e., X_2, X_5, X_6) should be allowed to correlate. Further, error in X_3 should be allowed to correlate with errors in both X_1 and X_2 because X_1 and X_2 are involved in the construction of X_3. These changes in the model are represented in Figure 1b. The only difference in the two models of Figure 1b is again the constraint on $\rho_{\xi_1\xi_2}$. Unconstrained, this correlation between the political-liberty factor and the popular-sovereignty one is estimated to be extremely high (.978). In fact, when the χ^2 values of the two models of Figure 1b are compared, the difference is nonsignificant ($\chi^2 = 0.59$, $p > .10$). This indicates that the two-factor solution no longer provides a superior fit compared to the simpler single-factor model, once correlated errors of measurement are introduced. Furthermore, the overall fit of both models in Figure 1b is good, judging from the nonsignificant χ^2 values.

Table 1 reveals that the covariances of the errors δ_3 with δ_1, and δ_5 with δ_2 are not significant for either model in Figure 1b. Therefore the models are reestimated constraining these error covariances to zero. Figure 1c shows the resulting models. The χ^2 values for both the bidimensional and unidimensional models are nonsignificant, indicating good fit for both models. In addition, the difference in χ^2 values is only .17, which is clearly not significant. Thus the hypothesis that the two factors are perfectly correlated—that the two factors can be represented as one

factor—is strongly supported. It is little wonder that this result follows when we note that even when $\rho_{\epsilon_1\epsilon_2}$ is unconstrained it is estimated to be .989 in Figure 1c.

AN EXTERNAL TEST

Using confirmatory factor analysis, we have found strong support for the hypothesis of a single dimension underlying the six indicators. In practice, however, simpler index-construction and regression techniques are used in most cross-national analyses. Suppose that we form a "political liberty index" from X_1, X_2, and X_3, and a "popular sovereignty index" from X_4, X_5, and X_6. If we substitute one index for the other in a regression analysis, will we obtain similar results? Such an exercise not only provides evidence of whether simpler techniques lead to the same conclusions as the more sophisticated ones, but also provides an "external" test of the dimensionality hypothesis and is of substantive interest as well. This test is external in that the relationship of the political democracy indicators to another variable is examined. In contrast, our confirmatory factor analysis was based on the relationships among the six indicators considered alone.

The example we choose is the relationship of political democracy to income inequality. The controversy in this area concerns whether political democracy reduces income inequality once economic development is controlled (e.g., Cutright, 1967; Jackman, 1975). Our income inequality measure is the Gini coefficient from Paukert (1973). Economic development is measured by another common proxy (cf. Grandjean, 1974), the logarithm of energy consumption per capita (Taylor and Hudson, 1971). We also include the square of this variable, in conformity with the specification of the development/inequality relationship suggested by the World Bank's researchers (see Weede, 1980).

The final explanatory variable is one of three indices. In the first equation Bollen's original political democracy index is used, the arithmetic average of the six indicators. The six indicators are scaled similarly, so we follow him in using equal weighting to combine them. In the second a political-liberty index is substituted, the average of the three political-liberty indicators X_1, X_2, and X_3. The third equation uses an index obtained by averaging the popular sovereignty indicators, X_4, X_5, and X_6. If these last two indices measure distinct dimensions, it is likely that they will have rather different effects on inequality. For example, popular sovereignty could have a substantial negative effect on inequality while political liberties might not. If so, the apparent effect of the combined index would obscure the separate effects of each dimension. On the other hand, if each index measures the same dimension, we would expect similar results for all of them.

Estimating these three equations using ordinary regression implicitly assumes no measurement error. Although here the reliabilities of the explanatory variables are likely to be high (e.g., Bollen, 1980), a more refined estimation procedure such as LISREL would explicitly take measurement error into account. Nevertheless, we use regression because we want to determine if this more familiar technique yields results consistent with our earlier LISREL analysis. In addition, the use of LISREL in models with polynomials of unmeasured variables has not yet been fully explored by statisticians. Our model includes a polynomial of economic development.[5]

In Table 2, all three indices of political democracy behave the same. None has an effect on inequality that would be significantly different from zero in a random sample of this size, all of the apparent effects are negatively signed, and they are

[5] In the absence of a well-developed theoretical model to guide research in this area, a full-information, maximum-likelihood (FIML) estimation approach like LISREL may be premature. FIML techniques are sensitive to specification error in any part of a system of equations. With the simplified structural model to be estimated, the more modest technique used here may be preferable. Furthermore, the sensitivity of LISREL-based estimates to relatively small sample sizes remains to be explored. Exclusion of nations with missing data on inequality or energy consumption left only 50 cases for this external test.

Table 2. Income Inequality (PAUGINI) Regressed on Each of Three Indices Measuring Political Democracy and Its Presumed Subdimensions (N = 50)

	Regression Coefficients (standard errors)		
	(1)	(2)	(3)
POLDEM	−.00048	—	—
	(.00067)		
POLLIB	—	−.00043	—
		(.00063)	
POPSOV	—	—	−.00043
			(.00064)
LNENPC60	.1223*	.1232*	.1215*
	(.0424)	(.0424)	(.0425)
$(LNENPC60)^2$	−.0110*	−.0111*	−.0109*
	(.0037)	(.0037)	(.0037)
INTERCEPT	.1796	.1726	.1810
	(.1229)	(.1205)	(.1246)
R^2	0.1838	0.1832	0.1827

NOTE: The following abbreviations for the variables are used:

PAUGINI = Paukert's (1973) Gini index of income inequality
POLDEM = Political Democracy index $= (X_1+X_2+X_3+X_4+X_5+X_6)/6$
POLLIB = Political Liberty index $= (X_1+X_2+X_3)/3$
POPSOV = Popular Sovereignty index $= (X_4+X_5+X_6)/3$
LNENPC60 = Natural log of energy consumption per capita, 1960
$(LNENPC60)^2$ = Square of the preceding variable
* Significant coefficient, $p<.01$ (see footnote 2 on significance test).

of virtually the same magnitude. Furthermore, the R^2 values for the three equations differ only in the third decimal place. These similar estimates regardless of the measure of political democracy again suggest that a single dimension does underlie the indicators.[6]

Our regression equations are based on a highly simplified specification of the democracy/inequality relationship. Still, because the same criticism could be made of other research in this controversial area, the substance of our results warrants at least brief comment. The negative regression coefficient for each of the three indices of political democracy is consistent with findings in some previous research (Rubinson and Quinlan, 1977; Weede, 1980). Unlike the studies just mentioned, we find a small and "nonsignificant" apparent effect for the 50 nations on which all data are available. The difference could mean that our specification omits a key suppressor variable, such as the state-strength variables emphasized by Rubinson and Quinlan. However, Weede reports that state strength has no effect on inequality when a quadratic term for economic development is included as we have done. Nor can our different results be attributed to our specification of the democracy/inequality relationship as fully recursive, because Rubinson and Quinlan suggest that ignoring reciprocal causation should inflate the apparent relationship. For reasons such as those treated at length by Jackman (1975) and Bollen (1980), we think the differences in results may stem largely from differences in the measures of political democracy or differences in the samples of cases analyzed. Our results support research indicating that, cross-nationally, the effect of political democracy on inequality is slight (e.g., Hewitt, 1977; Jackman, 1975). However, given the limited purposes of our analysis, we do not regard this finding as conclusive.

CONCLUSIONS

Theoretical considerations suggest two dimensions underlying political democracy: popular sovereignty and political liberties. However, both the regression results and our confirmatory factor

[6] Another external test yielded the same conclusion, without assuming perfect measurement. Instead of examining a possible consequence of democracy (equality), we looked at a possible cause, economic development. Simple structural-equation and measurement models were specified linking two development indicators (logged energy consumption and logged GNP per capita), the three political liberty indicators (X_1, X_2, X_3), and the popular sovereignty indicators (X_4, X_5, X_6). Using Jöreskog's methods, we tested the simultaneous constraints that (1) the effects of development on political liberty and on popular sovereignty are identical, and (2) the latter are perfectly intercorrelated. The χ^2 difference between the constrained and unconstrained solutions (1.43 with 2 degrees of freedom) would not be significant at the .05 level in a random sample, supporting the unidimensional solution. This support is less convincing than our other tests, however, because neither solution fits the data well, with χ^2 values of 26.31 and 24.88 suggesting underspecification. Also, the reservations in note 5 concerning FIML estimation of highly simplified models apply here as well.

analysis support its unidimensionality. Indeed, despite a bidimensional initial model, our search for the best fit arrived at the identical unidimensional final solution obtained by Bollen (1980) from a different starting point.

Is a bidimensional model wrong? It is tempting to answer an unequivocal yes, but to do so would neglect three important considerations. First, although we have not found two empirically distinguishable dimensions at the macro level, this does not rule out such a finding at the micro level. That is, individuals' attitudes toward political liberties and popular sovereignty may yet be distinct (see Weil, 1980).

Second, it must be remembered that our analysis rests on a cross-section of nations. Although cross-sectional variation in popular sovereignty and political liberties is unidimensional, it is possible that longitudinal data for one or more nations may reveal a bidimensional pattern. Bureaucratization viewed longitudinally appears more nearly unidimensional than cross-sectional studies of bureaucracy suggest (compare Meyer, 1972, with Pugh et al., 1968). The reverse could be true for democracy versus democratization.

Finally, we cannot say that a bidimensional model is "wrong" because this would neglect the important difference between theoretical and empirical dimensions. For some purposes one may still wish to distinguish the two dimensions in theoretical discussions. Yet any theory in which that distinction is made must also account for the almost perfect cross-national covariation between political liberty and popular sovereignty.

This difference between empirical and theoretical dimensions may lie at the heart of numerous debates about the "real" dimensionality of key sociological constructs (cf. Gold, 1964). There is much to be gained in these debates from research that makes explicit the assumptions underlying its operational definitions. Sociology needs to devote more attention to specifying measurement models along with structural models, even when the rigorous methods of Jöreskog for estimating such combined models are not applicable to the data at hand.

REFERENCES

Blalock, Hubert M., Jr.
1963 "Correlated independent variables: the problem of multicollinearity." Social Forces 42:233–7.
1979 "The presidential address: measurement and conceptualization problems: the major obstacle to integrating theory and research." American Sociological Review 44:881–94.

Bollen, Kenneth A.
1979 "Political democracy and the timing of development." American Sociological Review 44:572–87.
1980 "Issues in the comparative measurement of political democracy." American Sociological Review 45:370–90.

Bonjean, Charles M. and M. Grimes
1970 "Bureaucracy and alienation: a dimensional approach." Social Forces 49:363–73.

Cutright, Phillips
1963 "National political development." American Sociological Review 28:253–64.
1967 "Inequality: a cross-national analysis." American Sociological Review 32:562–78.

Dahl, Robert
1956 A Preface to Democratic Theory. Chicago: University of Chicago Press.

Downs, Anthony
1957 An Economic Theory of Democracy. New York: Harper.

Frisbie, W. Parker
1975 "Measuring the degree of bureaucratization at the societal level." Social Forces 53:563–73.

Gold, David
1964 "Criticism of an empirical assessment of the concept of bureaucracy: on conceptual independence and empirical independence." American Journal of Sociology 70:223–6.

Grandjean, Burke D.
1974 "The division of labor, technology, and education: cross-national evidence." Social Science Quarterly 55:297–309.

Hall, Richard H.
1963 "The concept of bureaucracy: an empirical assessment." American Journal of Sociology 69:32–40.

Hewitt, Christopher
1977 "The effect of political democracy and social democracy on equality in industrial societies: a cross-national comparison." American Sociological Review 42:450–64.
1979 "Reply to Stack." American Sociological Review 44:171–2.

Jackman, Robert W.
1975 Politics and Social Equality: A Comparative Analysis. New York: Wiley.
1980 "The impact of outliers on income inequality." American Sociological Review 45:344–7.

Jöreskog, Karl G.
1978 "Structural analysis of covariance and correlation matrices." Psychometrika 43:443–77.

Jöreskog, Karl G. and Dag Sörbom
1978 LISREL: Analysis of Linear Structural

Relationships by the Method of Maximum Likelihood. Version IV. Chicago: National Educational Resources.

Kohn, Melvin
1976 "Occupational structure and alienation." American Journal of Sociology 82:111–30.

Lenski, Gerhard
1966 Power and Privilege. New York: McGraw-Hill.

Lerner, Daniel
1958 The Passing of Traditional Society. Glencoe: Free Press.

Lipset, Seymour M.
1963 Political Man. Garden City: Anchor Books.

Meyer, Marshall W.
1972 "Size and the structure of organizations: a causal analysis." American Sociological Review 37:434–41.

Paukert, Felix
1973 "Income distribution at different levels of development." International Labor Review 108:97–125.

Pugh, Derek S., David J. Hickson, C. R. Hinings, and C. Turner
1968 "Dimensions of organizational structure." Administrative Science Quarterly 13:65–105.

Rubinson, Richard and D. Quinlan
1977 "Democracy and social inequality: a reanalysis." American Sociological Review 42:611–23.

Seeman, Melvin
1959 "On the meaning of alienation." American Sociological Review 24:783–91.

Stack, Steven
1979 "Internal political organization and the world economy of income inequality." American Sociological Review 43:271–2.
1979 "The effects of political participation and socialist party strength on the degree of income inequality." American Sociological Review 44:168–71.

Taylor, Charles L. and M. C. Hudson
1971 World Handbook of Political and Social Indicators II. Inter-Unniversity Consortium for Political and Social Research. Ann Arbor: University of Michigan.

Weede, Erich
1980 "Beyond misspecification in sociological analyses of income inequality." American Sociological Review 45:497–501.

Weil, Frederick D.
1980 "Tolerance of free speech in the United States and West Germany, 1970–1979: an analysis of public opinion survey data." Paper presented at the annual meeting of the American Sociological Association, New York City.

Wilensky, Harold L.
1967 Organizational Intelligence. New York: Basic Books.
1975 The Welfare State and Equality. Berkeley: University of California.

POLITICAL INVOLVEMENT AND ATTITUDE STRUCTURE IN THE GENERAL PUBLIC*

CHARLES M. JUDD
University of Colorado
JON A. KROSNICK
University of Michigan
MICHAEL A. MILBURN
University of Massachusetts–Boston

Differences in attitude structure as a function of political involvement are examined using a structural equation procedure with latent variables. In addition, criticisms of an earlier Judd and Milburn (1980) paper are examined. Findings suggest that the less involved are less likely to express political attitudes and that when they do their expressions contain more residual variance or error. At the same time, the attitudes of both the involved and uninvolved show evidence of a single underlying ideology that is remarkably stable in all groups. In arriving at these conclusions, a number of points are made concerning the procedures to be followed in comparing latent variable models between groups.

During the past twenty years sociologists and political scientists have studied the ideological organization of political attitudes in the general public. More specifically, they have examined differences in the degree of ideological organization between groups that differ in levels of political activity, interest in politics, and political sophistication (Achen, 1975; Converse, 1964; Converse and Markus, 1979; Judd and Milburn, 1980; Nie, Verba, and Petrocik, 1976). Philip Converse (1964) originally proposed that the organization of political attitudes differs between these groups in three ways: the degree to which attitudes reflect an underlying ideological predisposition, the degree to which attitudes are consistent with each other, and the degree to which attitudes are stable. The original research supporting the notion that attitude organization differs between groups reported correlations between attitudes (to examine consistency and stability) and responses to open-ended questions (to examine the presence of ideological thought).

A number of methodological problems in such procedures have recently been pointed out (Achen, 1975; Judd and Milburn, 1980). First, it is unclear whether responses to open-ended questions can be used to determine whether closed-ended attitude questions are ideologically based. A better procedure would be to look for an underlying ideological factor or construct that explains covariance in responses to the closed-ended questions. Second, comparisons of correlations between samples are affected by between-sample differences in the variances of the variables involved (Barton and Parsons, 1977). A better procedure would involve the comparison of unstandardized regression or structural coefficients (Blalock, 1967). Finally, two sorts of stabilities are confounded in correlations between the same attitude at two time points: ideological stability and the stability of residual or issue-specific variation. Judd and Milburn (1980) discussed these three problems at length and employed the technique of linear structural modeling in an attempt to overcome them. We examined the organization of five attitudes at three points in time in two samples that differed in educational attainment.

* Direct all correspondence to Charles M. Judd, Department of Psychology, University of Colorado, Boulder, Colorado 80309.

This research was supported by grants to the first author from the National Science Foundation (BNS 8005737) and The Henry A. Murray Center of Radcliffe College. The advice of David A. Kenny and an anonymous reviewer are gratefully acknowledged.

A number of criticisms have been raised about the Judd and Milburn (1980) approach (Converse, 1980a; Converse, 1980b). First, it has been suggested that our requirement of complete data on all fifteen attitude items (5 issues by 3 times) for inclusion in the sample had a large and differential impact upon the two samples. It is reasonable to assume that more respondents were deleted from the low-education group than from the high-education group due to missing data. Assuming that individuals with missing data have less organized attitude systems, it seems reasonable that the low-education group would have exhibited less stability and consistency had individuals with missing data somehow been included.

Second, it has been suggested that the amount of residual variance in the indicators of the Judd and Milburn (1980) model is an important indicator of attitude structure. The residual variance in an indicator reflects the degree to which variance in the measured attitude is due to factors other than the time-specific ideology and the issue-specific construct. In unstandardized models, a larger residual variance does not necessarily imply less impact of the ideological factor or of the issue-specific construct, but it does show that there is a greater amount of error in the observed attitude. The estimates of residual variance presented by Judd and Milburn (1980) are greater on the average for the low-education group than for the high-education group. We did not test the significance of this difference, however.

Finally, it has been suggested that Judd and Milburn's (1980) use of education to define the samples does not adequately address the hypothesis. Converse (1964:229) suggested originally that the samples should be defined according to "political interest." He later stated that measures of political involvement empirically discriminate between political "elites" and the general public better than does education (Converse, 1979, 1980a).

One purpose of this paper is to assess the validity of these criticisms of the Judd and Milburn (1980) paper. We explore the missing data problem, test the magnitude of residuals, and examine stability and consistency in groups that differ on political involvement as well as on education. Like Judd and Milburn (1980) we estimate the unstandardized coefficients of a latent variable structural model to overcome the problems in the original correlational procedures. We use the same data set as Judd and Milburn (1980): the 1972-1974-1976 election panel sample collected by the Center for Political Studies at the University of Michigan.

METHOD

Samples

It is now well established that there exist multiple dimensions of political involvement (Verba and Nie, 1972). Because we wish to compare groups that differ in involvement, it was necessary to identify the dimensions of involvement for the present sample. Hence, eleven involvement variables from the 1976 wave of the panel survey were factor analyzed, using a principle factoring procedure with iterated communalities. The eleven variables included measures of interest in the election campaign, self-reported efforts to persuade others, whether the respondent had signed petitions, whether he or she had written to newspapers or governmental officials, reported interest in following national affairs, the interviewer's judgment of the respondent's political knowledge, level of education, and a number of other potential indicators of political involvement. Three major factors accounted for the majority of variance in the unrotated factor solution, and hence three factors were rotated to enhance interpretation, using an oblique (Oblimin) rotation procedure.

The three involvement factors that emerged were readily interpretable. Loading highly on the first were campaign interest, national affairs interest, and whether the respondent reported trying to persuade others' electoral choices. This factor represents general interest and campaign activity, corresponding well with Verba and Nie's (1972) campaign activity factor. Loading highly on the second factor were the variables that assessed whether the respondent had written letters to political officials and/or newspapers. This factor seemed quite close to Verba

and Nie's (1972) contacting dimension. Loading on the third and final factor were respondent's education and the interviewer's judgment of respondent's political information. Thus, this factor seemed to be a general education or political information dimension.

In this paper we examine differences between groups that are high and low on each of these dimensions of involvement. To define the groups, we split the sample on three different variables, one loading highly on each of the three involvement factors. The three variables are expressed interest in the presidential campaign, whether or not the respondent reported writing to a public official, and education. To better distinguish extreme groups, we omitted those who fell in the middle of the three distributions. To further isolate extreme individuals on the first two dimensions, we selected respondents who were relatively stable in their level of interest and action during the four years of the panel study.

If we pay no attention to whether the attitude questions were answered and split the full panel sample of 1320, 308 respondents indicated in both 1972 and 1976 that they were "very much" interested in the presidential campaign. Of the remaining 1012 respondents, 597 reported that they were very interested in neither of the two years. The first group is defined as the high-interest group and the second as the low-interest group. We found that 264 respondents indicated in both 1972 and 1976 that they had written a public official on some political issue and 659 respondents did not indicate that they had written a public official either year. We defined the first sample as the high-write group and the second as the low-write group. We found that 257 respondents had attended four years of college or more by 1976 and 841 respondents had not attended college at all by 1976. The first group was defined as the high-education group, and the second was defined as the low-education group.

Measures

Judd and Milburn (1980) studied responses to five attitude questions at each of three time points. In order to conduct the present analysis with the largest possible samples, we specified a simpler model than the one used by Judd and Milburn (1980). The linear structural model can be estimated with as few as three different attitudes measured at two time points. Hence, in the present analysis, we examined attitudes in only 1972 and 1976, choosing the three attitude issues that resulted in the largest samples with complete data.

The attitude item on which there was by far the most missing data was the liberal/conservative self-rating scale used by Judd and Milburn (1980). Hence, it was omitted in the present analysis. As we would expect, many more individuals in the low interest, write, and education groups had missing data on this item than in the high groups. The large amount of missing data on this item, especially in the low groups, supports Converse's (1964) notion that the general public, particularly the less politically involved members of that general public, have a great deal of trouble expressing an ideological position.

In addition to deleting the liberal/conservative item, we did not use the measure of respondents' attitude about the rights of those accused of committing crimes, both because it had a substantial amount of missing data and because Judd and Milburn (1980) found it to be less consistent with the underlying ideology than were the remaining three indicators.

The three attitude measures that we did use are the following:

(1) Some people think achieving racial integration of schools is so important that it justifies busing children to schools out of their own neighborhoods. Others think letting children go to their neighborhood schools is so important that they oppose busing. Where would you place yourself on this scale (1 = Bus to achieve integration; 7 = Keep children in neighborhood schools)?

(2) Some people feel that the government in Washington should see to it that every person has a job and a good standard of living. Others think that the government should just let each person get ahead on his own. Where would you place yourself on this scale (1 = Government see to job and good standard of living; 7 = Government let each person get ahead on his own)?

(3) Some people feel that the government in Washington should make every possible effort to improve the social and economic position of blacks and other minority groups. Others feel that the government should not make any special effort to help minorities because they should help themselves. Where would you place yourself on this scale (1 = Government should help minority groups; 7 = Minority groups should help themselves)?

Table 1 presents the sample sizes for each of the six involvement groups. In the first column of Table 1, the sizes of the six groups are presented when we do not require any attitude responses. In the second column, we present the six sample sizes that result if we adopt Judd and Milburn's (1980) strategy of requiring complete data on five items at three time points. Below the sample sizes in this column are the percentages of the full samples in the six involvement groups with complete data. The third column presents the sample sizes when complete data are required on the three items used in this analysis at only two time points. Also presented in this column are the percentages of the full samples in the six groups having complete data on the six attitude questions. The models estimated in this paper used the samples in this third column.

It should be clear from this table that our samples represent a substantial improvement over those used by Judd and Milburn (1980). Nevertheless, we still must delete a substantial number of cases, particularly in the low-involvement groups, when complete data are required on only six attitude items. The disproportionate amount of missing data in the low-involvement groups supports Converse's (1980b) notion that the politically uninvolved are less likely to express political attitudes.

The Model

As we have said, the model used in this paper is a simplified version of Judd and Milburn's (1980) structural model. It is one of a set of models for panel data that has been suggested by Wheaton, Muthén, Alwin, and Summers (1977).

The model is defined by two structural equations:

$$X_{ij} = \alpha_{ij}T_j + \beta_{ij}I_i + U_{ij} \qquad (1)$$
$$T_2 = \gamma T_1 + \zeta \qquad (2)$$

In the first equation, called the measurement model, X_{ij} is a specific attitude of the three (i = 1 to 3) measured in either 1972 or 1976 (j = 1,2). T_j is the single time-specific latent variable that is assumed to explain all the covariation between the X_{ij} at any one point in time, and I_i is an item-specific construct which explains covariance between the same attitudes across time. For conceptual purposes, the T_j can be thought of as an underlying ideology. The α_{ij} is a loading coefficient of X_{ij} on T_j, and β_{ij} is a loading coefficient of X_{ij} on I_i. U_{ij} is residual variation in X_{ij} uncorrelated with T_j and I_i. These residuals (U_{ij}) are assumed to be uncorrelated with one another.

The second equation defines the structural effects of the latent ideology in

Table 1. Sample Sizes of Involvement Groups with and without Complete Data

	Total Sample	Judd & Milburn (1980) Complete Data (5 questions × 3 waves)	Present Model Complete Data (3 questions × 2 waves)
High Interest	308	163 (52.9%)	234 (76.0%)
Low Interest	597	178 (29.8%)	315 (52.8%)
High Write	264	158 (59.8%)	209 (79.2%)
Low Write	659	180 (27.3%)	355 (53.9%)
High Education	257	160 (62.3%)	205 (79.8%)
Low Education	841	226 (26.9%)	439 (52.2%)

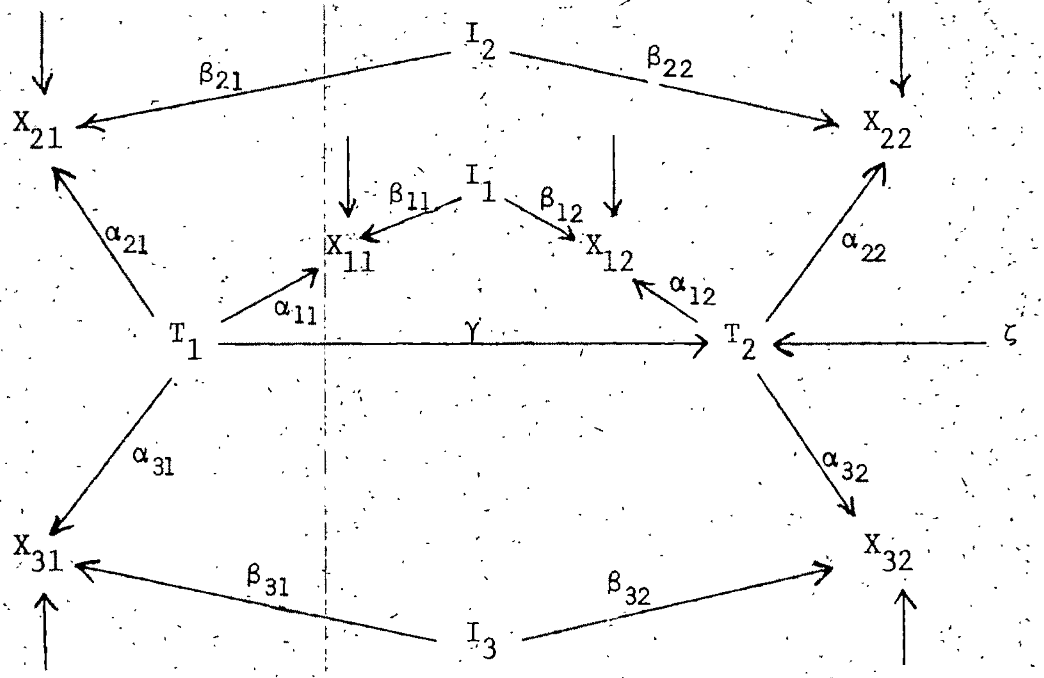


Figure 1. Diagram of the Measurement and Structural Models

1972 (T_1) on the latent ideology in 1976 (T_2). The stability coefficient is γ, and ζ is disturbance variance in T_2, uncorrelated with T_1. The measurement and structural models are diagrammed in Figure 1.

Estimation

The unknown parameters of the model were estimated using the LISREL procedure (Version IV; Jöreskog and Sörbom, 1978). This procedure generates maximum likelihood parameter estimates under the assumption of multivariate normality. The following unknown parameters of the model were estimated:

(a) the α_{ij} loading coefficients;
(b) the stability coefficient, γ;
(c) The U_{ij} disturbance variances;
(d) the ζ disturbance variance; and
(e) the variances of the latent variables ($\sigma^2_{T_1}$, $\sigma^2_{T_2}$, $\sigma^2_{I_1}$, $\sigma^2_{I_2}$, $\sigma^2_{I_3}$).

In order to estimate these parameters, the model must be identified; that is, there must be sufficient information in the variance-covariance matrix. In the present model, two kinds of restrictions or constraints must be put on the model in order for it to be identified. First, to estimate the variances of the latent constructs, one structural coefficient from each to an X_{ij} must be constrained at unity (Kenny, 1979). Second, since the same attitude is only measured twice, to estimate the variances of the item-specific factors (T_i), both structural coefficients from any I_i must be constrained at unity (i.e., $\beta_{i1} = \beta_{i2} = 1$).[1]

Assuming that the parameters of the model are identified, LISREL estimates them from the sample variance-covariance matrices. In the present case, the parameters of the model are overidentified and, hence, a goodness-of-fit test can be conducted to determine whether the hypothesized model is consistent with the data. For the present model the resulting χ^2 has 5 degrees of freedom.

In addition, the parameters of the model can be estimated simultaneously in multiple groups, and certain parameters can be constrained to be equal between groups.

[1] Instead of estimating the variances if I_i, they could have been fixed at some value and the β_{ij} coefficients could have been estimated, assuming that $\beta_{i1} = \beta_{i2}$. Such a model, however, is equivalent to the one we estimated, both structurally and substantively. A single structural parameter can be estimated for each I_i and it makes no difference whether it is the variance of I_i or β_{ij}.

Table 2. Maximum Likelihood Parameter Estimates; No Between-Group Constraints

	High Interest	Low Interest	High Write	Low Write	High Education	Low Education
α_{11}	.943	.537	1.011	.600	1.150	.543
α_{21}	1.007	.580	1.008	.775	.954	.823
α_{31}	1.000[a]	1.000[a]	1.000[a]	1.000[a]	1.000[a]	1.000[a]
α_{12}	.852	.472	.841	.584	1.038	.534
α_{22}	.861	.766	.981	.844	1.005	.937
α_{32}	1.000[a]	1.000[a]	1.000[a]	1.000[a]	1.000[a]	1.000[a]
γ	1.018	.855	.936	.990	.881	.889
ζ	.121	.768	.232	.445	.309	.416
$\sigma^2_{T_1}$	1.623	1.887	1.440	1.601	1.443	1.776
$\sigma^2_{T_2}$	1.801	2.148	1.492	2.014	1.429	1.820
$\sigma^2_{I_1}$	.664	.624	.606	.659	.840	.427
$\sigma^2_{I_2}$	.423	.477	.110	.406	.071	.484
$\sigma^2_{I_3}$	.189	.036	.494	−.100[b]	.255	.108
$\sigma^2_{U_{11}}$	1.165	.941	1.003	1.088	1.227	.977
$\sigma^2_{U_{21}}$	1.709	2.247	1.589	2.549	1.405	2.525
$\sigma^2_{U_{31}}$	1.648	1.381	1.092	2.062	.882	1.985
$\sigma^2_{U_{12}}$	.856	1.162	.915	1.122	.748	1.245
$\sigma^2_{U_{22}}$	1.559	1.346	1.372	1.606	1.008	1.503
$\sigma^2_{U_{32}}$	1.443	1.190	1.109	1.746	.913	1.657
$\chi^2(10)$	6.699		11.875		14.187	
p	.754		.294		.165	

[a] These coefficients were constrained at unity.

[b] This negative item-specific variance is nonsignificant. It indicates a nonsignificant negative relationship between X_{31} and X_{32} when ideology is controlled.

Such simultaneous estimates also yield goodness-of-fit tests that can be used to determine whether a model with between-group equality constraints is less consistent with the sample data than a model without such constraints.

RESULTS

As discussed in the previous section, the pattern of missing data is consistent with Converse's (1964) hypothesis that the less involved have difficulty expressing political attitudes. While many respondents have missing data on the attitude items, especially in the low-involvement groups, we can still examine whether attitude structures differ by political involvement among the majority in each group from whom we do have complete data.

When discussing estimation, we said that one structural coefficient from each of the latent ideologies (T_j) must be constrained at unity to derive their variances. In essence this defines the metric of the latent construct to be the same as the variable whose loading has been constrained. Judd and Milburn (1980) had a theoretical rationale for their choice of metric for the latent ideology. We constrained the loading of the liberal/conservative item, since the ideology was assumed to represent an underlying liberal/conservative ideological predisposition. In the present case, we have no theoretical rationale for the choice of metric of the latent ideology. Hence, we randomly chose to fix the loadings of the government aid to minorities items.

Table 2 presents the maximum likelihood parameter estimates from the six involvement groups. Also included in this Table are the χ^2 goodness-of-fit statistics for the simultaneous fit of the models in the high- and low-involvement groups with no between-group constraints. Regardless of how we define involvement, the hypothesized model is consistent with the data in the high- and low-involvement groups.[2]

The first step in examining whether the

[2] These goodness-of-fit statistics test whether the departures of the sample data from the variance/covariance matrices predicted by the model are significant. If they are not, and thus the χ^2 is not significant, the model is said to be consistent with the data (Kenny, 1979).

attitude structures differ between high- and low-involvement groups is to determine whether the measurement models are equivalent in the high and low groups. This involves constraining the α_{ij} coefficients to be invariant across involvement level and then examining whether the resulting models fit significantly less well than the unconstrained models. The difference between the χ^2s for the unconstrained and constrained models is itself a χ^2 and can be used to determine whether the constrained model fits less well. For the high- and low-interest groups, the difference $\chi^2(4) = 14.577$, $p < .01$; for the high- and low-write groups, the difference $\chi^2(4) = 7.415$; $p = .11$; and for the education groups, the difference $\chi^2(4) = 24.271$; $p < .001$. Hence, for interest and education, the α_{ij} coefficients are significantly different in the high- and low-involvement groups. For the write groups, the difference approaches significance. In the case of all three dimensions of involvement, then, the conclusion seems to be that the α_{ij} coefficients, the loading coefficients of the specific attitudes on the latent ideology, differ by level of involvement.

It is tempting to compare the α_{ij} coefficients of Table 2 between involvement levels and conclude that specific attitudes are less ideological in the low than in the high groups, since the α_{ij} coefficients are without exception smaller in the low-involvement groups. Such comparisons, however, must be made relative to the attitude whose loading has been constrained at unity. In other words, the unequal α_{ij} coefficients in Table 2 permit the conclusion that, relative to the aid-to-minorities issue, the busing and guaranteed-job issues are less ideological in the low-involvement than in the high-involvement groups. If another X_{ij} had been used to define the metric of the latent ideology, very different α_{ij} coefficients result. For instance, if the loading coefficients for the busing attitude are constrained at unity, the remaining α_{ij} coefficients are all larger in the low-involvement groups than in the high groups. In essence, then, the unequal loading coefficients permit us only to examine differences between involvement levels in the relative impact of ideology across issues. For the low groups, the busing issue is consistently the least ideological, while the aid-to-minorities issue is consistently the most. For the high groups, this ordering of issues is reversed: busing is the most ideological and aid to minorities the least. In the absence of a theoretical rationale for defining the metric of the latent ideology, comparisons of the overall impact of ideology across issues cannot be made between groups.

Examination of the coefficients of Table 2 suggests that the latent ideological factor is remarkably stable over the four-year interval in all six groups. All of the stability coefficients are greater than .85. These coefficients can be interpreted just as unstandardized regression coefficients are interpreted: for every one unit difference in the 1972 ideology we can expect .85 or greater units difference in 1976.

Is the latent ideology less stable in the low-involvement groups than in the high groups? Inspection of the coefficients of Table 2 suggests they are not. In the write groups, the stability coefficient is larger in the low group than in the high group. In the education groups, the stability coefficients are essentially equal. Only in the interest groups does there seem to be less stability among the less involved. However, such comparisons are also contingent upon the issue used to define the metric of the latent ideology so long as equal measurement models cannot be assumed between groups. With equal measurement models, the results of comparisons of stability coefficients are invariant across different metrics of the latent ideological factor. However, when the measurement models differ between groups, as we know they do in the present case, stability comparisons are contingent upon the metric used for the latent construct.

In essence, then, we can adopt one of three strategies in comparing stability coefficients between latent constructs in unstandardized models: (1) If we have a theoretical rationale for defining the metric of the latent factor, we can test for differences in stability under that metric (this is the strategy used by Judd and Milburn, 1980); (2) If the measurement model

is invariant across groups, we can compare stability coefficients between groups no matter which indicator is used to define the metric of the latent factors; (3) In the absence of a theoretical rationale for defining the metric, and when measurement models are known to differ between groups, comparisons of stability coefficients must be conducted defining the factor metrics all possible ways, i.e., using each indicator to fix the factor metric. If stability comparisons between groups yield essentially the same result regardless of the factor metric, we can have some confidence in the conclusions.

In the present situation, the third strategy must be adopted since we have shown unequal measurement models across involvement levels and since no theoretical rationale exists for the choice of the factor metric. Hence, for each of the three involvement dimensions, three different tests of equal stability in the high and low groups were computed, defining the metric of the factors using each of the three attitude items. The χ^2 tests of these nine comparisons are presented in Table 3. Each of the nine difference χ^2s in this table has a single degree of freedom. None of them approach significance. Hence, there is no evidence of unequal stability in high- and low-involvement groups, regardless of the dimension used to define involvement and regardless of the attitude item used to determine the metric of the latent ideology.

Table 3. Chi Squares Testing Equal Stability

	Item which Fixes Metric		
	Busing Issue	Guaranteed Job Issue	Aid to Minorities Issue
Interest Groups	.281	.253	.004
Write Groups	1.097	1.652	1.037
Education Groups	1.290	.770	.107

NOTE: All χ^2s have a single degree of freedom. All are nonsignificant.

The residual variances to the indicators ($\sigma^2_{U_{ij}}$) do not depend upon the metric used to define the factors. Hence, comparisons of the amount of residual error in the indicators yield the same result regardless of the factor metric. In Table 2, differences in the residuals are consistent in direction for both the write and education dimensions. As Converse (1980a) suggested, there seems to be more residual error in the low-involvement groups than in the high ones. This difference is not, however, apparent in the interest groups. Tests of equality of residual variances between high- and low-involvement groups reveal that: (1) in the write groups, residual variances are consistently and significantly larger in the low-involvement group than in the high, $\chi^2(6) = 21.14$, $p < .005$; (2) in the education groups, the residual variances differ significantly, $\chi^2(6) = 41.34$, $p < .001$. All but one of them are larger in the low group than in the high group. (3) In the interest groups, residual variances do not differ significantly by involvement level; $\chi^2(6) = 7.45$; $p > .25$.[3]

DISCUSSION

Judd and Milburn (1980) have developed an approach for examining differences in attitude structures that has distinct advantages over the original techniques used by Converse (1964) and others. Nevertheless, a number of criticisms have been made of this approach. In this paper, we have tried to evaluate those criticisms while keeping the advantages of Judd and Milburn's (1980) approach. Our results suggest that the issue of interest, differences in attitude structure, is an exceedingly complex one.

The following conclusions summarize

[3] One further apparent difference in the models between involvement groups deserves comment. The disturbances to the 1976 latent ideology, ζ, reported in Table 2, are higher in the low than in the high involvement groups. The magnitude of these disturbances, however, depends upon the metrics of the latent constructs. We tested for differences in ζ between groups under all three metrics. For the interest groups, ζ is higher in the low than in the high group under all three metrics, significantly so in two of the three cases. For the write groups, ζ is nonsignificantly higher in the low than in the high group under all three metrics. For the education groups, ζ is nonsignificantly higher in the low than in the high group under two of the three metrics. Under the third metric the high group has a nonsignificantly higher ζ. The general trend in these differences suggests that while ideology is equally stable between involvement groups, there is a tendency towards greater time-specific variation in the latent ideologies of the low involvement groups.

the results that have been presented in this paper:

(a) The pattern of missing data on the attitude questions supports the notion that those who are relatively uninvolved politically tend to have difficulty expressing political positions. This is particularly true when the attitude question asks the respondent to locate him- or herself along a fairly abstract ideological continuum.

(b) Among those who have complete data on the three issues used in the model we have examined, attitude structures differ between more and less involved groups in some ways but not in others. Different issues are ideologically derived to differing degrees in the high- and low-involvement groups. Some issues seem relatively more ideological for those with high involvement, while other issues seem to be relatively more ideological for those who are low on involvement. Thus, consistent with Judd and Milburn's (1980) findings, the role of ideology seems to depend jointly on the nature of the issue and the respondent's level of involvement. No matter how we define the latent ideological construct, there seems to be no evidence in our data for greater ideological stability among those who are more politically involved. At the same time, however, the attitude measures in the less involved groups seem to contain more error of measurement than in the high-involvement groups, so long as we use either education or political contacting to define political involvement.

(c) Generally, the results we have reported, concerning both the pattern of missing data and the attitude structures that emerge among those having complete data, do not depend upon the various ways of defining political involvement. The only exception to this conclusion concerns the fact that the attitude variables in the low-interest group do not have larger residuals than in the high-interest group, although differences in residuals are apparent when political involvement is defined in the other two ways.

These results support Converse's (1964) notions that there are differences between the more and less involved in the political attitudes they express. The less involved are more reluctant to express attitudes, and, when they do, those expressions are less reliable, in the sense that they reflect more measurement error. At the same time, however, many of Judd and Milburn's (1980) conclusions are also supported by our results. Among those who do express political attitudes, the presence of an underlying political ideology and the stability of such an ideology do not seem to depend upon the level of political involvement.

In arriving at these conclusions, we have made a number of methodological points concerning procedures for comparing latent structures. These methodological points are implicit in the statistical literature developing the factor-comparison techniques we have used (Jöreskog and Sörbom, 1979). Nevertheless, they are worth reiterating here.

First, comparisons between groups should generally be made with unstandardized coefficients, since such coefficients are relatively unaffected by differences in variability between groups. Second, the results of comparisons of structural relations among latent constructs depend upon the measurement model used to define those constructs. If measurement models are invariant across groups, then comparisons yield identical results regardless of the metric used to derive the variances of the latent constructs. If measurement models are not invariant across groups, then a theoretical rationale must exist for defining the metrics of the constructs. If such a rationale does not exist, then comparisons of structural coefficients between groups should proceed only if different metrics yield consistent results—that is, only if the results of the comparisons are consistent regardless of the metric used. Finally, comparisons of factor loading coefficients between groups must also be done relative to the metric of the factor. All of these points suggest that the researcher must be quite cautious in comparing covariance structures between groups.

While caution is appropriate, such comparisons can nevertheless yield results that are theoretically interesting. In the present case, these comparisons suggest a complex picture concerning the re-

lationship between political involvement and the organization of political attitudes. The less involved seem less able to articulate political attitudes. In addition, when they do, there is more variation in them that is predictable neither from an underlying ideology nor from the same attitudes at other time points. At the same time, however, the ideological base that is reflected in the various attitudes seems to be as reliable and stable in the low-involvement groups as in the high-involvement groups.

REFERENCES

Achen, Christopher H.
1975 "Mass political attitudes and survey response." American Political Science Review 69:1218–31.

Barton, Allen H. and R. Wayne Parsons
1977 "Measuring belief system structure." Public Opinion Quarterly 41:159–80.

Blalock, H. M.
1967 "Path coefficients versus regression coefficients." American Journal of Sociology 72:675–6.

Converse, Philip E.
1964 "The nature of belief systems in mass publics." Pp. 75–169 in David E. Apter (ed.), Ideology and Discontent. New York Free Press.
1979 "Comment." Journal of Politics 41:1182–4.
1980a "Comment: rejoinder to Judd and Milburn." American Sociological Review 45:644–6.
1980b "Recent thoughts on attitude constraint and stability." American Psychological Association annual meeting, Montreal.

Converse, Philip E. and Gregory B. Markus
1979 "Plus ca change . . .: the new CPS election study panel." American Political Science Review 73:32–49.

Jöreskog, Karl G. and Dag Sörbom
1978 LISREL: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood. Chicago: National Education Resources, Inc.
1979 Advances in Factor Analysis and Structural Equation Models. Cambridge, Ma.: Abt Books.

Judd, Charles M. and Michael A. Milburn
1980 "The structure of attitude systems in the general public: comparisons of a structural equation model." American Sociological Review 45:627–43.

Kenny, David A.
1979 Correlation and Causality. New York: Wiley.

Nie, Norman H., Sidney Verba, and John R. Petrocik
1976 The Changing American Voter. Cambridge: Harvard University Press.

Verba, S. and N. H. Nie
1972 Participation in America. New York: Harper and Row.

Wheaton, Blair, Bengt Muthen, Duane F. Alwin, and Gene F. Summers
1977 "Assessing reliability and stability in panel models." Pp. 84–136 in D. R. Heise (ed.), Sociological Methodology 1977. San Francisco: Jossey-Bass.

COMMENTS

NEW METHODS LEAD TO FAMILIAR RESULTS*

(COMMENT ON JUDD AND MILBURN, *ASR*, AUGUST 1980)

STEVEN S. MARTIN
University of Michigan

Judd and Milburn have presented a potentially important contribution to the debate about differences in the structure and stability of attitude systems between elite publics and mass publics. Their empirical work brings the welcome introduction of a structural equation approach to a field that has relied largely on correlational analysis. From their analysis they conclude, contrary to Philip Converse's work (1964, 1970, 1975), that a highly educated elite and a less educated nonelite both show relatively equal levels of attitude constraint and both show relatively equal levels of stability in the underlying ideological structure. These conclusions, however, are not supported by a more careful analysis of the authors' data. In fact, given the sample, the data provide some of the strongest published evidence for Converse's theories.

I make two major criticisms: that the sample selected for analysis loads the odds against finding support for Converse's conclusions and that Judd and Milburn's own data can be construed to support rather than challenge Converse's theories of attitude constraint and ideological stability.

Judd and Milburn analyze data for 464 out of 1,320 respondents in the CPS national election panel study (1972, 1974, 1976). This is 13% of the representative sample selected in 1972 and 35% of the complete panel. This smaller sample of respondents answered five scalar attitude questions asked in each of the three waves of the panel.[1] The authors break down these 464 into three subsamples based on the respondents' educational levels in 1976: the nonelite, who had not attended college (44%); the midelite, who had attended some college (25%); and the elite, who had completed four or more years of college (31%). This educational distribution contrasts sharply with the 1976 educational levels for the complete panel: 64% had not attended college, 17% had some college, and 18% had four or more years of college. The 65% mortality from the complete panel sample to the smaller sample employed by Judd and Milburn is far from random.

The explanation for the loss of cases lies in the wording of the five questions used in the analysis. All questions were asked as reported in the original paper with the addition that each question ended with the phrase, "or haven't you thought much about this?" The coding of "haven't thought about this" responses as missing data reduces the N from a potential of 1,011 complete data cases (an average of 1,281 per item) to 464. However, the "haven't thought about this" response is offered in the question and is chosen at some time by most of the panel respondents. So Judd and Milburn, in selecting respondents who have a scale position on all five issues at three points in time over four years, have created a very politically elite sample. This is suggested by the upward shift in the educational marginals noted above. Regressing an index of "haven't thought about this" responses on education reveals lower levels of education to be associated with a tendency to respond "haven't thought about this." And it is likely that those less educated respondents left in the study are the more politically interested and savvy of the less educated.

The use of education to separate elite versus nonelite publics comes from its convenience and from the fact that others, including Converse, have used it in the past.[2] In the present case its usefulness is minimized by the selection of a sample that is politically elite to begin with and then stratified by education. Under such restrictions it is not surprising that Judd and Milburn report that both elite and nonelite groups "show remarkable stability" in their attitudes and underlying ideological predispositions.

* Direct all correspondence to Steven S. Martin, Institute for Social Research, University of Michigan, Ann Arbor, MI 48109.

I wish to thank James House, William Mason, Duane F. Alwin, Philip Converse, Marie Crane, Marie Klatt, Cynthia Robbins, and an anonymous reviewer for their help. This research was supported in part by NSF Grant SOC-78-17407.

[1] There are actually 511 complete cases on the five variables, but the authors filtered on a sixth variable which they subsequently excluded from their analysis.

[2] In a recent comment (1979), Converse argues forcefully against the use of education as a criterion for discriminating between elite and mass publics.

What is surprising is that those model estimates can be construed to support the very premises they are supposed to contravene: the elite display more attitude constraint and more predictable ideological stability than the nonelite. I employ the same samples and largely the same numbers as those reported by Judd and Milburn,[3] though some results for the mid-elite group are also included. It appears that the authors neglected to report some data damaging to their thesis and made serious errors in statistical interpretation of LISREL lambda coefficients and in substantive interpretation of the political attitudes and factors.

Judd and Milburn estimated the unknown parameters in the following general model for the structure of attitude systems:

$$X_{ij} = \Sigma_k \alpha_{ijk} T_{jk} + \beta_{ij} I_i + U_{ij} \quad (1)$$
$$T_2 = \gamma_{21} T_1 + V_2 \quad (2)$$
$$T_3 = \gamma_{31} T_1 + \gamma_{32} T_2 + V_3 \quad (3)$$

Equation 1 is the measurement model: the effects of unmeasured constructs upon the observed variables. Equations 2 and 3 represent the structural model among the T_j factors. (See Judd and Milburn, 1980:633–4 for the definitions of the model terms.) The general model was estimated for each of the elite and nonelite groups simultaneously, with a χ^2 test performed to evaluate the fit of the model.

Judd and Milburn did not report on the mid-elite group, "in order to clearly discriminate between our elite and nonelite samples." The assumption is that differences should be most apparent in the two extreme educational samples. When they find a good fit for a single common factor structure for the extreme groups using a two-sample χ^2 test, they imply that the mid-elite behave similarly. The fit of the model with all three samples of education and no between-sample constraints, however, is not as good (χ^2 (216) = 263.23, p = .016). A test on the mid-elite sample alone finds a less than satisfactory fit: $\chi^2(72)$ = 97.14, p = .026.

This is a disturbing result. Judd and Milburn emphasize that the fit for each of the elite and nonelite groups does not deviate significantly from the predicted model. One would expect, given the authors' thesis, that the mid-elite sample would behave like the other two samples. In fact, statistically it should be easier to fit the model to the mid-elite because of its smaller sample size. There are at least three possible interpretations for this finding. First, with 48 free parameters per sample and samples of 143 (elite), 118 (mid), and 203 (nonelite) respondents, one simply cannot put much reliance on the estimates and certainly cannot generalize from them. The small number of observations per parameter means the estimates and tests are not very robust and are easily influenced by relatively small fluctuations in the data. Second, the elite and nonelite groups behave differently from those in between, perhaps because these particular variables are more relevant to the ideological distribution of those with high or low levels of education than to those in between. A third alternative interpretation, that the fit for the mid-elite group is good enough to consider it comparable to the other two samples, will be considered in detail later.

Even if one accepts Judd and Milburn's hypothesis of a common latent structure in each of the elite and the nonelite samples, their further discussion of attitude constraint and stability seriously misreads the results of the LISREL analysis. They begin, correctly, with the measurement model to test for an invariant factor pattern by equating the α_{ij} coefficients between the two samples. This procedure follows the pattern of analysis outlined by Jöreskog (1971) and substantively provides a test for equality of attitude constraints in the two samples. The analysis reveals a difference in χ^2 with a poor fit, showing that α_{ij} coefficients are not identical across the two samples. Trying to salvage some support for equality (de facto if not de jure) of the α_{ij} coefficients, Judd and Milburn offer a speculative argument: although the α_{ij} coefficients are not equal across samples, "it still may be the case that they have α_{ij} coefficients of relatively equal magnitudes across all variables" (p. 638). They conclude that this is the case, based on a comparison of the α_{ij} coefficients: 5 out of 12 coefficients are greater for the elite sample, 7 out of 12 are greater for the nonelite sample; the average α_{ij} for the elite is 1.17, for the nonelite, 1.31.

To get these covariance estimates of the α_{ij} coefficients, Judd and Milburn fix a parameter for each of the factors, establishing a common but arbitrary metric for factors across groups. But, in avoiding the problem of comparing standardized coefficients between samples, they neglect new concerns raised by setting the metric. A direct comparison of the α_{ij} coefficients is simply not warranted. Although the coefficients are unstandardized, they are ratios whose metric is set in terms of a variable, X_{6j} (liberal-conservative), that has different values and variances in each sample: that is, no α_{ij} has a real value; only the ratio α_{ij}/α_{6j} is defined. The number of larger and smaller α_{ij}'s between samples depends on the particular metric es-

[3] The nonelite sample used here is the same as in the earlier paper. There is a small disparity in the elite sample reported here (making no appreciable difference in the analysis) likely due to my working with the later release of the panel data.

tablished for the factor when a variable is chosen to be fixed. For example, if α_{3j} (aid to minorities) is fixed at 1.0 rather than α_{5j} (liberal-conservative) fixed by Judd and Milburn, 8 out of 12 coefficients are greater for the elite sample and the average α_{ij} for the elite is 0.94, and for the nonelite, 0.75. One cannot compare models for the two samples by looking at the sizes of the α_{ij} coefficients. The same reasoning prevents one from examining the coefficients of stability, the β_{ij}'s, and then drawing conclusions regarding their relative magnitudes across samples.

Judd and Milburn state that choosing a variable "to set the metric of a latent construct is arbitrary" (p. 635). It is apparent, however, that for the comparison of coefficients the choice has far from arbitrary consequences. They also assert that there may be a theoretical rationale for fixing the liberal-conservative coefficient: it refers to the general ideological predisposition rather than to a specific issue. Such an assumption requires some support or at least discussion on their part. On face value it appears a weak assumption because most of the missing data discussed above comes on this variable and because, as Converse and Markus (1979:42, 43) point out, the nature of this variable leads to erroneously high estimates of individual-level stability.

Having concluded that the measurement model is not equivalent in both elite and nonelite samples, and without establishing a theoretical rationale for the metric imposed on the parameter estimates, Judd and Milburn cannot compare the *magnitude* of coefficients. Although one cannot make direct comparisons of magnitude between samples, one can compare the proportion of variance in each item attributable to the latent constructs.

The variance of items in terms of factors in the model, given that the items and factors have unit variances, can be decomposed as follows:

$$S^2_{xi} = 1.0 = \alpha^2_{iT} + \beta^2_{ii} + \psi_i$$

where α^2_{iT} is the variance accounted for by the T_j factor, β^2_{iT} the variance accounted for by the I_i factor and ψ_i the unexplained variance.[4] It is possible within each sample to consider the α^2_{iT} a measure of the proportion of constraint of the X_{ij} into the factor T_j, and consider the β^2_{ii} a measure of the proportion of stability of the X_{ij} in the item-specific factor I_i, remembering that T_j and I_i are defined as orthogonal to each other. Table 1 presents these decompositions of proportion of variance within each sample for all 15 items.

The obvious finding revealed in Table 1 is that the "constraint" attributed to the ideological factor T_j is a larger proportion for the elite sample than for the nonelite on all 15 items. The elite sample consistently displays a higher proportion of individual item variance constrained through the underlying ideological construct than does the nonelite sample. This is in accord with Converse's contention and counter to the argument of Judd and Milburn. An examination of the variance attributable to the item-specific factor produces a mixed pattern: 7 items have a larger proportion of variance for the elite sample, 8 items for the nonelite sample. It is important to note, though, that the item stability estimated through LISREL is not the same as the item-specific stability Converse discusses. For the LISREL model, stability is orthogonal to constraint and represents the variance in a question that is not attributable to the ideological factor but can be attributed to the factor of the same item asked at three points in time. For Converse, stability and constraint are overlapping constructs and part of what Converse would consider stability is encompassed in the ideological factor as well as the item-specific factor. Therefore, the evidence available from a reexamination of the measurement model supports Converse's notion on attitude constraint but describes a related but different notion of attitude stability.

Turning next to the structural model, Judd and Milburn compare the ideological stability of elite and nonelite. Again a model specifying equality of coefficients across the two samples (here the γ_{ij} coefficients) produces a poor fit to the data. Their further examination concentrates on differences in the γ_{31} coefficient: the direct effect from 1972 to 1976. They find, "in the elite sample, the effect of T_1 on T_3 seems to be nearly totally mediated by T_2" (γ_{31} elite = 0.05), "while in the nonelite sample there exists a substantial unmediated effect of T_1 on T_3" (γ_{31} nonelite = 0.49). Judd and Milburn conclude that the significant difference in γ_{31} across the two samples is evidence that, "the nonelite population shows more long term stability in their underlying ideology than the elite population" (p. 640).

The existence of the large γ_{31} coefficient for the nonelite sample, however, points to the volatility in their ideological position. A first order Markovian model of stability (or change) posits that one's ideological position at time 2 is a function of position at time 1, and one's ideological position at time 3 is a function of one's position at time 2. A significant effect of position at time 1 on time 3 unmitigated by

[4] In terms of factor analysis, $\alpha^2_{iT} + \beta^2_{iT}$ is referred to as the communality of the item and ψ_i is the variance unique to the item.

Table 1. Decomposition of Item Variance from LISREL Estimates

	Elite			Nonelite		
	Constraint	Stability	Unexplained	Constraint	Stability	Unexplained
BUSING72	.521	.234	.245	.307	.172	.521
RIGHTS72	.282	.247	.471	.219	.314	.467
MINORITIES72	.570	.019	.411	.392	.126	.482
GUARANTEED72	.494	.025	.481	.386	.775	−.161[a]
LIBCON72	.502	.216	.282	.349	.288	.363
BUSING74	.571	.291	.138	.318	.200	.482
RIGHTS74	.260	.327	.413	.159	.441	.400
MINORITIES74	.600	.024	.376	.410	.260	.330
GUARANTEED74	.576	.137	.287	.379	.046	.575
LIBCON74	.542	.320	.138	.240	.464	.296
BUSING76	.553	.180	.267	.240	.533	.227
RIGHTS76	.133	.549	.318	.107	.231	.662
MINORITIES76	.495	.486	.019	.439	.132	.429
GUARANTEED76	.529	.048	.423	.439	.021	.540
LIBCON76	.481	.226	.293	.440	.141	.419

[a] I have reported this anomaly as it was presented by Judd and Milburn. A negative error disturbance is a potentially serious problem. In these data this condition occurs in the variable GUARANTEED72 in each of the nonelite and mid-elite samples. Further examination, however, reveals that this term is not significant in either sample and is likely the result of sampling fluctuation. Called a Heywood case, the conventional practice is to fix the LY coefficient in LISREL at the estimate obtained with the negative variance, then to constrain the disturbance term to zero and reestimate the model.

position at time 2 is a form of instability in a group's underlying ideology. This is consistent with Wheaton et al.'s (1977) explanation of stability as a lag-1 system. For Judd and Milburn to argue otherwise requires an explanation for the substantial γ_{31} effect in the nonelite sample. Without explanation one must conclude that the elite sample demonstrates more stability in the ideological factor over time than does the nonelite sample.

Further support for the ideological stability of the elite being greater than that for the nonelite comes from an examination of the R^2's between the T_j factors, reported in Table 2. Both elite and nonelite samples show substantial factor stability, as evidenced by the variance explained in T_2 by T_1 and in the variance explained in T_3 by T_1 and T_2. However, the elite sample R^2 is larger at both times than the R^2 for the nonelite sample, particularly when one compares the R^2 for T_2 across samples. In sum, the structural model reexamination suggests that the attitudes of the elite are not as subject to an off-year election defocusing (T_2 is the 1974 off-year election) as are the attitudes of the nonelite. All of which again supports the contention of more constraint and stability for the elite attitudes, which remain salient even without the fanfare of a presidential election.

It is worth returning to the model estimates for the mid-elite group.[5] If on the whole, the attitude constraint and the stability of the factor structure for the mid-elite sample fall between those estimated for the elite and for the nonelite groups, then there is additional support for Converse's hypotheses. If on the other hand, the mid-elite estimates produce no discernible pattern, then the operationalization of elite in terms of educational levels does not provide for a test of Converse's hypotheses or of Judd and Milburn's alternative.

Table 2. R^2 for the T_j Factors

	Elite	Nonelite
T_2	.911	.706
T_3	.901	.867

Figure 1 and Table 3 provide the estimates for the mid-elite sample. Figure 1, analogous to Figures 2 and 3 in Judd and Milburn, presents the α_{ij} (constraint) coefficients and the γ_{ij} (factor stability) coefficients for the mid-elite model. Table 3 presents the decomposition of proportion of variance for both the α_{iT} and the β_{ii}.

[5] If one rejects the fit of the mid-elite model, an examination of parameter estimates is not meaningful. However, many researchers would consider $\chi^2 = 97$, df $= 72$ as an adequate fit, particularly if not seen in conjunction with the better fits for the elite and for the nonelite. If one accepts the fit for all three samples, they can be considered as sharing a common factor structure with more or less attitude constraint and factor stability from sample to sample. The mid-elite parameter analysis reported here then becomes relevant.

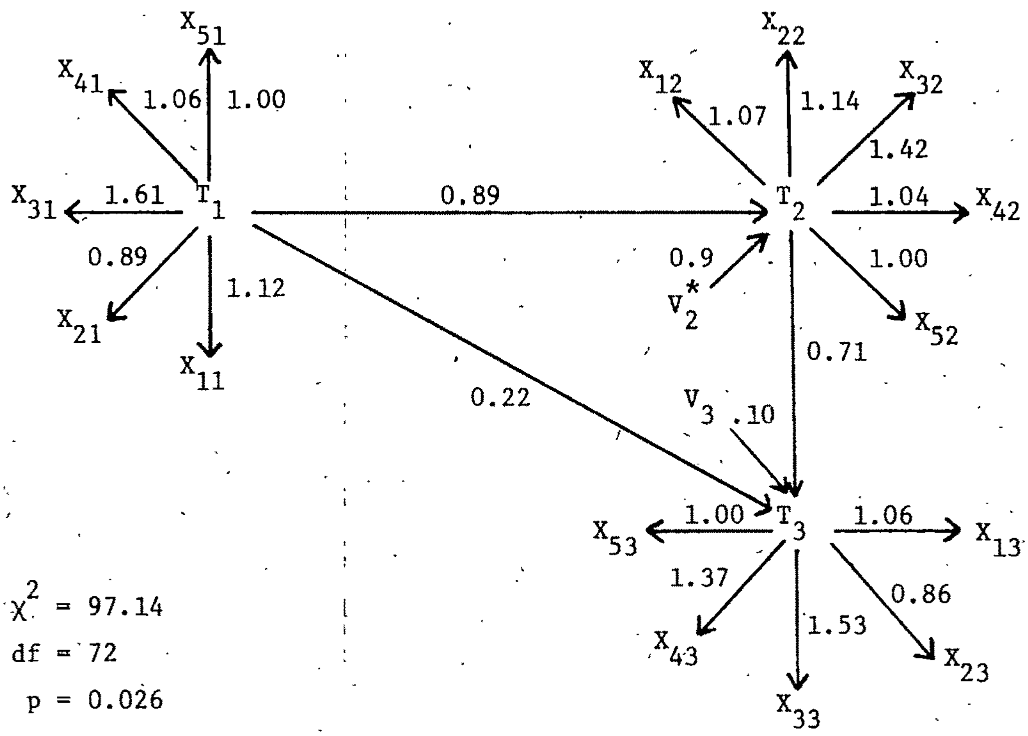


Figure 1. Maximum-Likelihood Parameter Estimates: Mid-elite Sample (N=117)
NOTE: Although Judd and Milburn refer to V_j as the variance of the disturbance to the latent endogenous construct, they report standard deviations in their analogous figures. The numbers reported here are variances.

Table 3. Decomposition of Item Variance from LISREL Estimates: Mid-Elite Sample

	Constraint	Stability	Unexplained
BUSING72	.361	.338	.301
RIGHTS72	.149	.304	.547
MINORITIES72	.622	.159	.219
GUARANTEED72	.300	.894	−.194
LIBCON72	.438	.245	.317
BUSING74	.357	.442	.201
RIGHTS74	.241	.558	.201
MINORITIES74	.533	.216	.251
GUARANTEED74	.346	.041	.613
LIBCON74	.352	.332	.316
BUSING76	.586	.395	.019
RIGHTS76	.130	.468	.402
MINORITIES76	.534	.134	.332
GUARANTEED76	.506	.024	.470
LIBCON76	.362	.313	.325

The estimates of attitude constraint in Table 3 for the mid-elite sample fall between those of the elite and nonelite samples for eight of the fifteen items, four estimates are lower than those for the nonelite group, and three estimates are higher than those for the elite group. The evidence is mixed, but on the whole it offers some support to Converse. An examination of the γ_{ij} coefficients in Figure 1 is more clear-cut. Here the coefficient of concern again is γ_{31}, which for the mid-elite group is 0.22 and falls neatly between that for the elite sample (0.05) and for the nonelite sample (0.49).

Though the small-sample mid-elite estimates must be treated with caution, if one is willing to accept the overall fit of the mid-elite model as roughly equivalent to that of the elite and nonelite models, the evidence tilts toward Converse's position: the mid-elite on the whole do fall between the elite and the nonelite in attitude constraint and ideological stability. If one finds the fit of the mid-elite group unacceptable, neither Converse's hypotheses nor alternatives are testable with these data.

The weaknesses in Judd and Milburn's paper are substantial. From sample selection through statistical analysis to substantive interpretation, they attempt to force a conclusion the data will not support. They have made an important contribution in separating distinct components of constraint and stability in attitude structures. Empirically and conceptually

the two components overlap in Converse's work, and the structural equation approach provides an empirical means to parcel out separate estimates of them. But in their attempt to disconfirm Converse's hypotheses, they have turned a helpful analytic tool into blinders, theoretically and empirically. In any case, a definitive resolution of the elite-mass debate cannot be made with these data. At the very least researchers need to operationalize a criterion that emphasizes the "political" in distinguishing political elites from political nonelites.

REFERENCES

Converse, Philip E.
1964 "The nature of belief systems in mass publics." Pp. 206–31 in David E. Apter (ed.), Ideology and Discontent. New York: Free Press.
1970 "Attitudes and non-attitudes: continuation of a dialogue." Pp. 168–89 in E. R. Tufte (ed.), The Quantitative Analysis of Social Problems. Reading, Mass.: Addison-Wesley.
1975 "Public opinion and voting behavior." Pp. 75–169 in F. I. Greenstein and N. W. Polsby (eds.), Handbook of Political Science, Vol. 4. Reading, Mass.: Addison-Wesley.
1979 "Comment." The Journal of Politics 41:1182–4.
1980 "Comment." American Sociological Review 45:644–6.

Converse, Philip E. and Gregory B. Markus
1979 "Plus ça change . . .: the new CPS election study panel." The American Political Science Review 73:32–49.

Jöreskog, Karl G.
1971 "Simultaneous factor analysis in several populations." Psychometrika 36:409–26.

Jöreskog, Karl G. and Dag Sörbom
1978 LISREL IV. Chicago: National Educational Resources, Inc.

Judd, Charles M. and Michael A. Milburn
1980 "The structure of attitude systems in the general public: comparisons of a structural equation model." American Sociological Review 45:627–43.

Wheaton, Blair, Bengt Muthen, Duane F. Alwin, and Gene F. Summers
1977 "Assessing reliability and stability in panel models." Pp. 84–136 in D. R. Heise (ed.), Sociological Methodology 1977. San Francisco: Jossey-Bass.

Wray, J. Harry
1979 "Comment on interpretations of early research into belief systems." The Journal of Politics 41:1173–81.

INTERPRETING NEW METHODS IN ATTITUDE STRUCTURE RESEARCH*

(REPLY TO MARTIN)

MICHAEL A. MILBURN
University of Massachusetts-Boston

CHARLES M. JUDD
University of Colorado

There have been many interpretations of Philip Converse's hypothesis concerning the belief structures of the mass public. These versions can be ordered roughly along a continuum of extremity. At the most extreme, Converse's work has been understood as implying that the vast majority of the public has "nonattitudes" on the major issues that confront the country. The least extreme version of the hypothesis argues that the belief systems of the public differ at least in some predictable ways from the belief systems of political elites. Even critics of Converse's work, including ourselves, acknowledge the validity of this most modest form of the hypothesis. The debate concerns not if there are differences in belief systems but the magnitude and nature of those differences. We believe, despite Steven Martin's comment, that the evidence still supports our earlier argument (Judd and Milburn, 1980) that most members of the general public, regardless of their level of education or political involvement, show remarkable stability and consistency in their political attitudes.

To demonstrate this point, we address the four major criticisms that Martin has made of our paper: (1) that the composition of our samples biases our results in favor of high stability and consistency and that the use of education to discriminate groups is inappropriate; (2) that we failed to report results from a "mid-elite" group (results which he suggests are inconsistent with our conclusions); (3) that our choice of the liberal-conservative self-rating question as a metric for the latent variables is inappropriate and thus comparisons of factor loadings to infer differences in attitude constraint are incorrect; and (4) that our interpretation of the stability coefficients is inaccurate.

Sample Selection

Martin argues that our samples do not adequately represent the mass public because we

* Direct all correspondence to Charles M. Judd, Psychology Department, University of Colorado, Boulder, Colorado 80309.

We wish to thank Jon A. Krosnick for his helpful suggestions. This work was supported by a National Science Foundation grant (BNS 8005737) to the second author.

excluded those who indicated they had not thought much about the issues. Thus, he argues, we were dealing with a "nonelite" sample that at the very least professed political opinions. Certainly there is validity in this argument. Our samples are substantially reduced due to missing data. However, it is not at all clear that our results would be much different had we been able, somehow, to obtain responses from every individual in the sample.

Judd, Krosnick, and Milburn (1981) examine a two-wave model similar to our three-wave model, but they use a substantially smaller number of attitude items. By using fewer items at only two points in time, Judd et al. (1981) are able to test the model on nonelite samples that are twice the size of those used by Judd and Milburn. They reach the same conclusions about attitude stability and consistency as we did earlier. The data from both elite and nonelite groups are consistent with a single ideological factor model with high stability.

Martin also criticizes our use of education level to discriminate between elites and nonelites. Judd et al. (1981) show that our conclusions are generally valid when measures of political involvement, which are more consistent with Converse's theoretical notions, are used to discriminate between samples.

In general, criticisms of the samples we used are well-taken. However, the available evidence suggests that the use of more inclusive samples does little to alter our original conclusions.

Examination of the "Mid-Elite" Sample

In discussing results from a mid-elite sample, Martin suggests that our failure to present results from this group is misleading and that analyses of this group render the analyses we report questionable. We make three points in response to this argument.

First, we can quite honestly say that we excluded this mid-elite sample from our analysis for theoretical reasons rather than because the data were damaging to our thesis. We wanted to discriminate between extreme education groups. We excluded from our analysis those who were not extreme. It might be argued that for theoretical reasons we should have examined this middle group. But we did not. An inference that we chose not to publish their data because they were inconsistent with our conclusions is simply incorrect.

Second, we question Martin's suggestion that, since the model does not fit data from the mid-elite sample as well as it fits data from the elite and nonelite samples, our findings lack credibility. It is not clear to us why the lack of fit in one sample invalidates the fit of the model in the other two samples.

In spite of the lack of fit of the model for this mid-elite sample, Martin presents the results on their data and suggests that these results are consistent with Converse's hypothesis. For reasons we discuss below we have profound disagreements with his procedure at this point. But, even if we grant him his standardization, the results he reports are equivocal at best. He argues that the standardized loadings he presents should fall between the standardized loadings from the elite and nonelite samples. Eight of the fifteen do so; the remaining seven do not. Even if we grant Martin his procedure (which we do not), his results are hardly conclusive.

Assessment of Constraint

Martin criticizes the procedure by which we compared constraint in the elite and nonelite groups. We found that a model constraining the unstandardized loadings on the ideological factors to be equal across samples fits the data poorly. We then observed that the average loading in the unconstrained model for the nonelite sample is slightly higher than that for the elite sample. We argued that these results cannot be used to support the conclusion that the nonelite group exhibits less constraint than the elite group does.

Martin correctly points out that the choice of an indicator with which to establish a metric for a particular latent construct (by fixing the indicator's loading at 1.0) affects the relative magnitude of other indicators' loadings on that construct. Martin argues that it is always inappropriate to compare unstandardized loadings across samples because different metrics yield different results. We disagree. When one has a theoretical reason for selecting a particular variable with which to establish a metric, it is appropriate to make such comparisons. The variance of a latent construct is defined as the variance in the chosen indicator that is shared with the other indicators of the construct. Since we wished to define the latent constructs in our model as ideological positions along a liberal-conservative continuum, fixing the metric with an indicator explicitly measuring individuals on such a scale was quite reasonable. Comparisons of unstandardized loadings across groups under these circumstances seem perfectly appropriate.

The alternative procedure for examining constraint that Martin proposes is clearly inappropriate. It amounts to examining standardized parameter estimates. Such a procedure requires the assumption that the variances

of the attitudes in the two groups are equal. We simply are not willing to make this assumption. Comparisons between samples require an examination of unstandardized structural coefficients using an appropriate metric for the latent constructs. This is one of the fundamental advances we hope to have made in examining Converse's hypothesis.

Assessment of Stability

We finally come to Martin's discussion of what he calls our "most telling interpretive weakness:" our interpretation of the stability estimates for the latent ideology. Martin suggests that because there is a strong relationship between the 1976 ideological position of the nonelite sample and both their 1972 and 1974 positions, this is evidence of instability. On its face, this argument does not seem to make much sense. He goes on to argue that his interpretation is consistent with attitude stability as a lag-1 system. If we examine Judd and Milburn's results as a lag-1 system, we find that they support our conclusion of roughly equal stability in the two groups. In other words, a one unit change in T_1 is associated with equal changes in T_2 in both groups. Likewise, a one-unit change in T_2 is associated with equal changes in T_3 in both groups.

Once again using standardized parameter estimates, Martin continues his discussion of stability by examining the proportion of variance explained (R^2) over time in the latent ideologies of the two samples. He concludes that the nonelite sample shows off-year defocusing of their attitudes to a greater extent than does the elite sample. This conclusion is inconsistent with the conclusion reached when the more appropriate unstandardized structural coefficients are examined. If substantial off-year defocusing occurred for the nonelite, then we would expect to find low 1972 to 1974 stability, low 1974 to 1976 stability, and high 1972 to 1976 stability as the presidential elections bring the attitudes into focus. This is simply not the case. The two-year stability estimates are quite high for both elites and nonelites, and the four-year stability estimate is higher for the nonelite sample. To argue that the higher four-year stability coefficient in the nonelite sample means that their attitudes are less stable can be done only if one ignores the remarkably high two-year stabilities in both groups.

Conclusion

We believe that Martin's criticisms have done little damage to our general conclusions. We have shown that more inclusive samples do not alter the results significantly. Likewise, examining differences in structure as a function of political involvement does not affect our conclusion. The mid-elite group, which we failed to examine earlier, also shows high attitude constraint and stability, even though their data are somewhat less consistent with the model. Martin's critique of our metric for the latent ideology constructs seems inadequate. Our metric is a theoretically meaningful one that permits us to compare unstandardized structural coefficients. It seems to us that the use of such a metric is clearly preferable to Martin's insistence on standardization. Finally, if we accept the lag-1 stability model Martin proposes, our results still indicate equally high stability in the elite and nonelite groups.

There are undoubtedly some differences in the nature and organization of political attitudes between those who are politically sophisticated and those who are not. Nevertheless, our results demonstrate a remarkable amount of both attitude stability and consistency in all segments of the general public. We still think that attitude theorists should find comfort in these results.

REFERENCES

Judd, Charles M., Jon A. Krosnick, and Michael A. Milburn
1981 "Political involvement and attitude structure in the general public." American Sociological Review 46:000–000.

Judd, Charles M. and Michael A. Milburn
1980 "The structure of attitude systems in the general public: comparisons of a structural equation model." American Sociological Review 45:627–43.

"SEX AND AUTHORITY IN THE WORKPLACE": A REPLICATION AND CRITIQUE*

(COMMENT ON WOLF AND FLIGSTEIN, *ASR*, APRIL 1979)

WILLIAM P. BRIDGES
BAILA MILLER
University of Illinois at Chicago Circle

Wolf and Fligstein (1979:250) recently concluded that, in order to move into supervisory positions at the same rate as men, "First,

* Direct all correspondence to William P. Bridges, Department of Sociology, University of Illinois at Chicago Circle, Chicago, IL 60680.

women should be encouraged to improve their qualifications. But this is not enough to alleviate the sex gap in authority in the workplace. More important, steps must be taken to alter the behaviors and policies of employers. . ." We argue that this conclusion, while perhaps warranted in the long run, is unjustified given the data and procedures of the authors. Ignoring the theoretical issues raised by their uncritical adoption of the status-attainment paradigm, we are troubled by their choice of data and their interpretation of its statistical properties. In particular, we show that their findings are weakened by the facts that (1) attention is limited to those who attended and finished high school in Wisconsin; (2) attention is limited to a population of high-school graduates at midlife (around age 37); and, most importantly, (3) their interpretation of components in a regression decomposition analysis is arbitrary and the model is misspecified. By analyzing a different data set, we replicate a portion of their analysis and demonstrate empirically the consequences of each of these limitations.

The three dependent variables in Wolf and Fligstein's (1979) study are based on responses of men and women in the 1975 sample of the Wisconsin Study of Social and Psychological Factors in Socioeconomic Achievements to questions regarding aspects of authority in the workplace: (1) whether R has the "authority to hire and fire others;" (2) whether R can "influence or set the rate of pay received by others;" and (3) whether R can "supervise the work of others, that is what they produce or how much." The last of these items was asked in nearly identical form in the 1972 Institute for Social Research Quality of Employment Survey, although the two higher levels of authority went untapped. The specific wording of the supervision question was, "Do you supervise anyone as part of your job?"

The presence of this item opens up the possibility that the results from the Wisconsin, "midlife," high-school graduate sample can be compared to those from a nationally representative sample. This possibility is abetted by the presence of measures of the independent variables in the Quality of Employment Survey (QES) that are nearly identical to those used by Wolf and Fligstein. We believe that where the correspondence is inexact, the validity of the QES measures is, on balance, equal, if not superior, to those in the Wisconsin survey. The measures available in QES are as follows: (1) Education—an eight-point ordinal scale recoded in this analysis to exact year (i.e., high-school graduate equals twelve) or midpoints of appropriate intervals. In our replication section these scores have been reduced by a value of twelve to represent years after high school. (2) Tenure—an eight-point ordinal scale of time with current employer recoded into months. (3) Work experience—this was calculated by dividing respondents' answers to the question, "About how many years in total have you worked for pay since you were sixteen years old?" by the value of their age minus sixteen. This is less satisfactory than the measure constructed by Wolf and Fligstein but surpasses common measures based on age and schooling calculations and is calculated consistently for both sexes. (4) Marital status—coded one for currently married respondents, 0 otherwise. (5) Children—the presence of child-care responsibilities is measured by recoding the responses to a question asking how many children under age 16 there are, of whom R is a parent, in R's household, with all values greater than zero set equal to one. (6) Status—Duncan SEI score of respondent's current job. (7) Male and (8) Unlabeled are measures of the sex composition of R's current job, identical to those used by Wolf and Fligstein except that our recoding was carried out on percentages given in the QES data set which are taken from Occupation by Industry census tables rather than from Occupation census tables. Because of this greater detail, the QES measure probably has higher validity. The means and standard deviations of the variables used in the replication portion of the reanalysis (high-school graduates, aged 30–49) are shown in Table 1.[1]

While the purpose of this comment is not to carry out a detailed analysis of marginal frequencies, two observations can be made. First, the two samples are very nearly identical in terms of occupational characteristics. Both men and women have a slightly greater chance of being supervisors in the national sample (perhaps because they are slightly older), but the difference between men and women is nearly the same in the two samples—.227 nationally vs. .229 in the Wisconsin survey. In terms of so-called "human capital" factors, the national sample has more work experience and tenure and fewer children. Again, these are differences which could stem from the broader definition of midlife which is necessary with

[1] The means, standard deviations, and regression coefficients reported in Tables 1 through 4 have all been calculated applying the appropriate sample weights given in the QES data set. The number of cases and the standard errors have all been calculated assuming that the sample size is 1,496, the unweighted number of cases. This was accomplished by converting the sample weights to fractional values by multiplying them by the ratio of the unweighted total N to the weighted total N.

Table 1. Means and Standard Deviations from QES Data Compared to Wisconsin Data

	QES Sample				Wisconsin Sample			
	Male		Female		Male		Female	
Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Supervise Others	.641	(.481)	.414	(.495)	.607	(.484)	.378	(.485)
Education After HS	1.89	(2.22)	.904	(1.46)	1.90	(2.44)	1.10	(1.89)
Work Experience	.892	(.134)	.634	(.272)	.818	(.156)	.585	(.258)
Tenure	103.8	(84.6)	82.4	(73.6)	90.5	(60.5)	47.5	(53.7)
Married	.920	(.272)	.752	(.434)	.892	(.310)	.789	(.408)
Children	.744	(.437)	.573	(.497)	.870	(.336)	.811	(.391)
Status	51.2	(23.0)	45.2	(21.0)	50.8	(22.8)	46.1	(20.1)
Male	.608	(.489)	.064	(.245)	.633	(.482)	.062	(.241)
Unlabeled	.382	(.487)	.433	(.498)	.345	(.475)	.370	(.483)
N of Cases (Unwtd.)	209		109		3359		2254	

the smaller national sample, but they also could be associated with the fact that the population of Wisconsin tends to underrepresent nonwhites. Thus, despite the rather limited size of the QES sample in the required age and education ranges, it conforms rather well to empirical expectations.

To replicate the analysis of Wolf and Fligstein, we begin with the midlife, high-school educated, national sample and apply their "reduced-form" and unrestricted regression models. We present only the latter in Table 2, but refer to the reduced-form results where necessary. In the QES male equation the large negative effects associated with the male occupation and the unlabeled occupation should not be taken seriously inasmuch as they are almost surely the result of multicollinearity. (Note that only 1% of males are in female occupations.) When the unlabeled occupation variable is dropped from the equation, the male occupation variable has a positive regression coefficient equal to .080 with a standard error of .067.[2] Thus, it is quite similar to the results of Wolf and Fligstein. There are a number of other points of resemblance between the two samples. In both, the returns to occupational status for males are slightly under twice those of females, and the coefficients across samples are roughly equal. Furthermore, education for females has a large direct effect on their chances of becoming supervisors in both samples, while the direct effects of male education are rather small in both. In the reduced-form model not shown here, there is a significant effect of male education on supervisory status equal to .046, a figure quite close to that obtained in the Wisconsin data, .050. The magnitude of the R-squared values is the same across both data sets.

While these results tend to confirm those of Wolf and Fligstein, there are some major differences. Both work experience and tenure have larger effects for men than women in the national sample, lending support to the oft-repeated charge that women tend to become mired in "dead-end" jobs. The national sample also shows a large handicap faced by women with children—being responsible for one or more children under sixteen cuts chances of becoming a supervisor by 1/3. For some reason this pattern does not obtain for Wisconsin-educated women; instead of a negative effect for children, we find a positive effect for work experience. This is difficult to explain unless women who were educated in Wisconsin are more likely than the national sample to leave the labor force in response to childbearing. Some fragmentary census evidence bears out this conclusion, although it would be surprising if this mechanism accounted for all the discrepancy in the two samples.[3] In any event, the results from the national sample require a

[2] Dropping the unlabeled occupation variable does not substantially affect the regression coefficients for other variables in the equation.

[3] Thus, 1970 population census figures for Wisconsin and the U.S. show the following. Among Wisconsin women who were aged 16–44 in 1970 and who had no children under age 6 in 1965 and who were working in 1965, of the 58,799 who *did* have young children by 1970, 33% (or 19,446) continued to work. In comparison, of the 157,436 who did not have young children in 1970, 84.8% (or 133,524) continued to work. However, for the same group of women in the United States generally, of the 4.92 million who did have young children by 1970, 36% (1.78 million) continued to work, compared to the 14.6 million who did not have young children in 1970, of whom 80.6% (or 11.8 million) continued to work. Thus, the relative employment dropout rate for women with children is higher for the Wisconsin sample than for the national sample. Using the phi measure of association, the analogue of Pearson's r for dichotomies, the degree of association for the United States is −.42 compared to −.51 in Wisconsin.

Table 2. Replication of Wolf and Fligstein's (1979) Unrestricted Regression Model

	QES Sample				Wisconsin Sample			
	Male		Female		Male		Female	
Variable	B	S.E.	B	S.E.	B	S.E.	B	S.E.
Education	−.0002	(.019)	.068*	(.034)	.018	(.005)	.034*	(.006)
Work Experience	.302	(.260)	−.087	(.180)	.040	(.077)	.255*	(.052)
Tenure	.00075*	(.00038)	−.0009	(.0007)	.00022	(.00014)	.0007*	(.0002)
Marital Status	.055	(.133)	.139	(.107)	−.023	(.029)	−.0009	(.027)
Children	−.065	(.079)	−.338*	(.098)	.085*	(.027)	.010	(.033)
Status	.0076*	(.0018)	.0040	(.0024)	.0070*	(.0004)	.0040*	(.0005)
Male	−.506	(.322)	−.141	(.189)	.092	(.054)	.064	(.041)
Unlabeled	−.600	(.322)	.012	(.092)	.019	(.055)	−.055	(.021)
Intercept	.440		.398		−.045		−.041	
R-Squared (ADJ)	.120		.132		.144		.108	

* p<.05.

somewhat different interpretation than that suggested by Wolf and Fligstein's data. Thus, it is not that women's human capital is so much more highly rewarded than men's, net of occupational characteristics, but that their life circumstances are more directly constraining on their authority advancement. Whether this results from the behavior of women or from invidious actions of employers (or of husbands) cannot be ascertained without additional data.

Given these discrepancies, a relevant question is how well the results of Wolf and Fligstein (1979) generalize to a population defined more broadly in terms of age and education. In Table 3, means, standard deviations, and regression results are shown for all adult wage earners and salaried workers represented in the QES sample. Because the educational attainment restriction has been lifted in this sample, one adjustment is needed in the way variables

Table 3. Means, Standard Deviations, and Regression Results for QES Sample—All Ages, All Education Levels

	Men		Women	
Variable	Mean	S.D.	Mean	S.D.
Supervise Others	.4933	(.500)	.2900	(.454)
Education after HS	1.20	(1.85)	.717	(1.24)
Education before HS	11.12	(1.95)	11.22	(1.65)
Work Experience	.854	(.208)	.681	(.259)
Tenure	92.27	(91.98)	67.40	(77.70)
Married	.800	(.401)	.620	(.486)
Children	.514	(.500)	.397	(.490)
Status	39.62	(24.15)	39.04	(20.62)
Male	.637	(.481)	.065	(.247)
Unlabeled	.342	(.475)	.369	(.483)
N of Cases	618		414	

Regression Results

	Reduced-Form Model				Unrestricted Model			
	Men		Women		Men		Women	
Variable	B	S.E.	B	S.E.	B	S.E.	B	S.E.
Education after HS	.057	(.011)	.056*	(.019)	.003	(.012)	.028	(.020)
Education before HS	.019	(.010)	.012	(.014)	−.003	(.010)	−.008	(.015)
Work Experience	.077	(.101)	.035	(.088)	−.028	(.097)	.018	(.088)
Tenure	.0007*	(.00023)	.00038	(.0003)	.00046*	(.00022)	.00022	(.0003)
Married	.122*	(.056)	.044	(.048)	.051	(.056)	.036	(.047)
Children	.035	(.044)	−.039	(.048)	.030	(.042)	−.026	(.048)
Status	—	—	—	—	.0080*	(.0010)	.0049*	(.0012)
Male	—	—	—	—	.166	(.131)	.040	(.091)
Unlabeled	—	—	—	—	.133	(.132)	−.054	(.046)
Constant	−.036		.059		−.024		.140	
R-squared (ADJ)	.080		.022		.168		.055	

* p<.05.

Table 4. Regression Results from QES Data Set with Status by "Collar-Color" Interactions

	Men		Women	
Variable	B	S.E.	B	S.E.
Education after HS	.002	(.010)	.032	(.020)
Education before HS	−.003	(.013)	−.007	(.015)
Work Experience	−.033	(.098)	.017	(.087)
Tenure	.00048*	(.00022)	.00025	(.0003)
Married	.049	(.056)	.037	(.047)
Children	.034	(.042)	−.028	(.048)
Status (If Blue Collar)	.0065*	(.0018)	.0105*	(.0038)
Status (If White Collar)	.0054*	(.0020)	.0047*	(.0024)
White Collar Dummy	.182	(.132)	.086	(.149)
Male	.201	(.133)	.026	(.091)
Unlabeled	.145	(.132)	−.065	(.047)
Constant	−.026		.052	
R-Squared (ADJ)	.171		.056	

*$p<.05$.

are represented in the equation. To permit a comparison of education coefficients before and after the high-school-completion cutting point, the effect of education is represented as a linear spline function. This is accomplished by constructing two new variables. The coefficient of the first represents the effect of education up to and including a high-school diploma; the coefficient of the second represents the effect of education beyond high school (both coefficients are estimated subject to a continuity constraint).[4]

Education continues to have a larger effect for women than for men after the occupation variables are added to the model, although it is now insignificant for both sexes. However, despite the fact that returns to education beyond high school are present in the reduced-form models for both sexes, this finding cannot be generalized to workers having no more than a high-school education. In fact, for both sexes the effect of having no education beyond high school in the reduced-form models is three to four times less than the effect of post-high-school education. Another interesting facet of these results is that job tenure has a larger effect for men than women, and work experience is basically unimportant for either group, supporting the results obtained when the QES population is limited to those who completed high school. Furthermore, the impediments to a woman's becoming a supervisor created by children in the household are no longer apparent when the age and education restrictions are removed. The only remaining similarity to the initial results of Wolf and Fligstein is found in the fact that there is still a large disparity favoring men in the male and female returns to occupational status. Because this difference is crucial to Wolf and Fligstein's finding of employer discrimination in the distribution of authority, it must be examined more closely.

One possibility is that, here, again, a trend that cannot be generalized to the total universe appears in a limited portion of the universe of workers. Table 4 explores this possibility with regard to occupational status by allowing occupational status to have differing effects for white- and blue-collar workers of both sexes. In addition, to specify the model properly, a dummy variable is included which takes a value of one if the respondent is a white-collar worker and remains zero otherwise.

This slight respecification produces a drastic shift in the overall pattern of results. Among blue-collar workers, occupational status is more strongly related to the chances of becoming a supervisor for women than for men. Among white-collar workers occupational status has a slightly more positive effect for men than women and men do receive a somewhat larger "authority bonus" for crossing the white-collar line. However, these latter two patterns are overshadowed by the large blue-collar reversal. Thus, it is clear that different processes govern access to authority in different occupational strata. Furthermore, if the findings of Stolzenberg (1975) concerning variations in income attainment across specific occupations and Parcel (1979) concerning variation in income attainment across race groups can be taken as precedents, these white-

[4] The coding scheme used to produce these variables is described by Featherman and Hauser (1978: 261 ff.) and is a slight modification of a technique developed by Poirier (1975). The basic idea is to allow the slope of a given independent variable to vary in different segments of its range. This is achieved by replacing the vector of scores on the independent variable with a number of new vectors (as many as there are segments to be estimated). The regression coefficients which result from the technique used here can be interpreted as the actual slope in the given segment—i.e., before or after high school.

collar/blue-collar differences in authority attainment may be just the tip of the iceberg.

The decomposition of the male-female authority gap for the models in our Tables 2 and 4 appears in the upper panel of Table 5. The first decomposition illustrates the effects of broadening the sample to a national group, and the second, the effects of removing age and educational restrictions and of respecifying the prestige-authority relationship. However, both decompositions are very misleading, having been based on a regression analysis with categorical variables coded as dummy variables. With dummy-variable coding, decomposition results can change depending on which category of the regressor is suppressed in the coding scheme (see proof in Appendix). Duplicating Wolf and Fligstein, we suppressed the nonmarried category of the marital-status variable, the childless category of the children variable, and the female category of the occupation/sex variable. However, had we made other choices, the decomposition results would have changed.

A straightforward solution to this problem is the use of effect coding. When this is done, regression decomposition results are unaffected by the decision to omit one category rather than another. However, the choice between dummy-variable coding or effect coding cannot be an arbitrary one. The use of dummy-variable coding means that an arbitrary selection of one group as a comparison group can produce an equally arbitrary change in regression components (not coefficients) and in substantive interpretation. The use of effect coding is desirable because the arbitrary choice of which category is suppressed has no effect on statistical or substantive results.

The decompositions, using effect coding, of both QES models (see bottom panel of Table 5) can be compared to the results presented in Wolf and Fligstein's (1979:249) Table 5. While they found that the component with the largest positive value was the job-rates component, our restricted sample results show that the largest is now the human-capital/family-rates component. Furthermore, when a sample of a broader population is examined, the basis of the employer discrimination interpretation completely disappears because the job-rates component is now negative.

Apart from the size of the job-rates component, there is also the question of how it is to be interpreted. According to Wolf and Fligstein's logic, the job-rates component is unique be-

Table 5. Decomposition of Components of Male/Female Difference

A. With Dummy Variable Coding	Original Model (See Table 2)		Expanded Model (See Table 4)	
Component	Amount	Pct.	Amount	Pct.
Total	.2272	100.0%	.2035	100.0%
Human Capital/Family Total	.5095	224.2	.0579	28.4
HC/FAM Composition	−.0096	−4.2	.0285	14.0
HC/FAM Rates	.4173	183.7	.0371	18.2
HC/FAM Interaction	.1018	44.8	−.0077	−3.7
Job Characterisics Total	−.3243	−142.7	.2231	109.6
Job Char Composition	−.0536	−23.6	.0421	20.7
Job Char Rates	−.1251	−55.1	.1353	66.5
Job Char Interaction	−.1455	−64.0	.0457	22.4
Interaction	.0420	18.5	−.0775	−38.1

B. With Effect Coding	Original Model (See Table 2)		Expanded Model (See Table 4)	
Component	Amount	Pct.	Amount	Pct.
Total	.2272	100.0%	.2035	100.0%
Human Capital/Family Total	.4151	182.7	.0208	10.2
HC/FAM Composition	−.0094	−4.1	.0282	13.8
HC/FAM Rates	.3228	142.1	.0000	0.0
HC/FAM Interaction	.1017	44.8	−.0074	−3.6
Job Characteristics Total	.0014	0.6	.0464	22.8
Job Char Composition	−.0536	−23.6	.0414	20.3
Job Char Rates	.2006	88.3	−.0416	−20.4
Job Char Interaction	−.1455	−64.0	.0466	22.9
Interaction	−.1893	83.3	.1363	67.0

cause it, more than any other component, reflects only the discriminatory actions of employers. This logic depends on the assumption that SES and sex composition are adequate measures of the potential for a job involving supervisory responsibility. While there is little evidence for or against this proposition, the realization of that potential depends on the motivation and preferences of workers and the desires of coworkers, as well as on the discriminatory behavior of employers. One need look no further than a group of tenured professors in an academic department to realize that the desire to become a supervisor (department chair) on the part of each professor is as important in determining who gets the position as are the dean's biases. Thus, without relevant measures of employees' attitudes and employer's behavior, and in the absence of a convincing rationale for the discrimination interpretation of the job-rates component, the less said about having demonstrated the role of employer discrimination the better.[5]

To conclude, we believe that Wolf and Fligstein (1979) selected an important avenue of investigation but have assumed too much and measured too little. Not only are substantial portions of their findings inapplicable beyond a rather narrowly defined universe, but their techniques and analytic logic can be used to produce findings which directly contradict their own. We do support their realization that more data on employer behavior and worker attitudes are needed, and would add that policies for redistributing workplace authority between men and women can be developed only when an adequate specification of relevant processes has been established.

[5] Of course, differences between men and women which are created outside of the workplace may also be due to discrimination. The question is whether the authority gap is fabricated out of thin air by employers or whether it results from a more comprehensive social shaping process.

REFERENCES

Featherman, David and Robert Hauser
1978 Opportunity and Change. New York: Academic Press.

Parcel, Toby
1979 "Race, regional labor markets, and earnings." American Sociological Review 44:262–79.

Poirier, Dale
1975 "On the use of bilinear splines in economics." Journal of Econometrics 3:23–4.

Stolzenberg, Ross M.
1975 "Occupations, labor markets and the process of wage attainment." American Sociological Review 40:645–65.

Wolf, Wendy C., and Neil D. Fligstein
1979 "Sex and authority in the Workplace: the causes of sexual inequality." American Sociological Review 44:235–52.

APPENDIX

Proof of the equality of "rates" regression coefficient with the use of "effect" coding. This proof applies to the most basic situation—i.e., the two-subpopulation, two-level categorical predictor case. It could be generalized (tediously) to the two-subpopulation, multi-level categorical predictor case.

Basic model:

$$y_m = a_m + b_m x_m + c_m z_m + e_m \quad (1)$$

$$y_f = a_f + b_f x_f + c_f z_f + e_f \quad (2)$$

Where:

- $y_{m\ (or\ f)}$ = dependent variable in subpopulation m (or f)
- $a_{m\ (or\ f)}$ = intercept in subpopulation m (or f)
- $x_{m\ (or\ f)}$ = vector of codes representing two-level categorical predictor x in subpopulation m (or f)
- $b_{m\ (or\ f)}$ = regression coefficient for x_m (or x_f)
- $z_{m\ (or\ f)}$ = noncategorical predictor variable in subpopulation m (or f)
- $c_{m\ (or\ f)}$ = regression coefficient for z_m (or z_f)

Also defined are:

- N_{im} = Number of cases at level i (1, 2) of subpopulation m (or f)
- b_{1m} = Intercept for subpopulation m (or f) at level 1 of x
- b_{2m} = Intercept for subpopulation m (or f) at level 2 of x
- $b_{\cdot m}$ = "Unweighted" intercept for subpopulation m (or f)

WITH DUMMY VARIABLE CODING:

With category 1 "suppressed" (level 1=0)

In Equations 1 and 2:

$a_m = b_{1m}$ $\qquad a_f = b_{1f}$

$b_m = b_{2m} - b_{1m}$ $\qquad b_f = b_{2f} - b_{1f}$

$$\bar{x}_m = \frac{N_{2m}}{N_{1m} + N_{2m}} \qquad \bar{x}_f = \frac{N_{2f}}{N_{1f} + N_{2f}}$$

Rates component

$$RC_1 = ((b_{2m} - b_{1m}) - (b_{2f} - b_{1f})) \frac{N_{2f}}{N_{1f} + N_{2f}} + (c_m - c_f)\,\bar{z}_f$$

With category 2 "suppressed" (level 2=0)

In Equations 1 and 2:

$a_m = b_{2m}$ $\qquad a_f = b_{2f}$

$b_m = b_{1m} - b_{2m}$ $\qquad b_f = b_{1f} - b_{2f}$

$$\bar{x}_m = \frac{N_{1m}}{N_{1m} + N_{2m}} \qquad \bar{x}_f = \frac{N_{1f}}{N_{1f} + N_{2f}}$$

Rates component

$$RC_2 = ((b_{1m} - b_{2m}) - (b_{1f} - b_{2f})) \frac{N_{1f}}{N_{1f} + N_{2f}} + (c_m - c_f)\,\bar{z}_f$$

Comparing the two, let:

D_{1s} = slope difference with 1 "suppressed" = $((b_{2m} - b_{1m})-(b_{2f} - b_{1f}))$

D_{2s} = slope difference with 2 "suppressed" = $((b_{1m} - b_{2m})-(b_{1f} - b_{2f})) = -D_{1s}$

$k = (c_m - c_f)\,\bar{z}_f$

Then:

$$RC_1 = D_{1s}\frac{N_{2f}}{N_{1f}+N_{2f}} + k = (-D_{2s})\left(1 - \frac{N_{1f}}{N_{1f}+N_{2f}}\right) + k$$

$$= -D_{2s} + RC_2$$

Thus, in general the rates component differs depending on whether category 1 or category 2 is suppressed.

EFFECT CODING:

With category 1 "suppressed" (level 1= −1)

$a_m = b_{\cdot m}$ $\qquad a_f = b_{\cdot f}$

$b_m = b_{2m} - b_{\cdot m}$ $\qquad b_f = b_{2f} - b_{\cdot f}$

$$\bar{x}_m = \frac{N_{2m} - N_{1m}}{N_{1m} + N_{2m}} \qquad \bar{x}_f = \frac{N_{2f} - N_{1f}}{N_{1f} + N_{2f}}$$

Rates component

$$RC_1 = ((b_{2m} - b_{\cdot m}) - (b_{2f} - b_{\cdot f}))\frac{N_{2f} - N_{1f}}{N_{1f} + N_{2f}} + k \text{ (as above)}$$

With category 2 "suppressed" (level 2= −1)

$a_m = b_{\cdot m}$ $\qquad a_f = b_{\cdot f}$

$b_m = b_{1m} - b_{\cdot m}$ $\qquad b_f = b_{1f} - b_{\cdot f}$

$$\bar{x}_m = \frac{N_{1m} - N_{2m}}{N_{1m} + N_{2m}} \qquad \bar{x}_f = \frac{N_{1f} - N_{2f}}{N_{1f} + N_{2f}}$$

Rates component

$$RC_2 = ((b_{1m} - b_{\cdot m}) - (b_{1f} - b_{\cdot f}))\frac{N_{1f} - N_{2f}}{N_{1f} + N_{2f}} + k \text{ (as above)}$$

Because:

$$(b_{2m} - b_{\cdot m}) = -(b_{1m} - b_{\cdot m}) \text{ and}$$

$$(b_{2f} - b_{\cdot f}) = -(b_{1f} - b_{\cdot f}) \text{ and}$$

$$\frac{N_{2f} - N_{1f}}{N_{1f} + N_{2f}} = -\frac{N_{1f} - N_{2f}}{N_{1f} + N_{2f}} \text{ or } \bar{x}_{1f} = -\bar{x}_{2f}$$

Then:

$$RC_1 = D_{1s}\bar{x}_{1f} + k = (-D_{2s})(-\bar{x}_{2f}) + k \qquad \text{Where } D_{1s} \text{ and } D_{2s} \text{ and } k \text{ are as defined above}$$

$$= RC_2$$

Thus, in general rate components are equal, regardless of which categories are suppressed when effect coding is used. The expression used for the rates component is an algebraic simplification of that used by Wolf and Fligstein, but is equivalent.

RESPONSE TO BRIDGES AND MILLER*

NEIL FLIGSTEIN
University of Arizona and
National Opinion Research Center

MICHAEL SOBEL
Bureau of Labor Statistics

WENDY C. WOLF
University of Arizona and
National Commission for Employment Policy

We welcome the chance to engage in a substantive debate on the issues raised by our recent paper. Unfortunately, Bridges and Miller (1981) have added little to that debate. In this comment, we will directly confront the issues raised by Bridges and Miller (1981) that relate to our data and method. We are troubled that, on the whole, Bridges and Miller confirm our results while their tone and argument suggest that they have radically undermined our findings. We think that they have done little to confront or extend the theoretical issues we raise.

Before considering the explicit criticisms, we feel it necessary to respond to a comment about our "uncritical adoption of the status-attainment paradigm." Discussion of what is and is not status-attainment research is often trivial. The precursor to our piece (Wolf and Fligstein, 1979b) and this piece itself challenge the research result of status equality between the sexes and argue that using status as the measure of one's position in the hierarchy of work obscures our understanding of stratification in the workplace. As such, we view our work as part of the ongoing redefinition of the stratification area. We are, whether or not Bridges and Miller think so, active participants in challenging the conventional wisdom on measuring and conceptualizing key dimensions of work.

Bridges and Miller have three points: (1) our sample and variable measures are problematic, (2) our model is misspecified, and (3) the regression decomposition we perform is arbitrary. We answer each in turn.

Bridges and Miller argue that our sample is restricted, making generalization a bit difficult. Furthermore, they argue that their replication with the QES (and equally good if not better measures) gives a more accurate representation of the process of acquiring power in the workplace. We will show that when they estimate our models using their expanded sample, their parameters and decompositions are quite similar to ours, suggesting that our sample problems must not be so severe. In terms of the sample, we are quite explicit about the problems:

> Using this data set means that there are no individuals with less than 12 years of education included in the sample. The results cannot be generalized to non-high school graduates. Furthermore, we are investigating the distribution of authority at midlife (around age 37) and our results do not address the issue of the distribution of authority in the work setting for the total working population or for one cohort earlier or later in their life course. [Wolf and Fligstein, 1979a:239]

We chose to use the Wisconsin data and not the Quality of Employment Data for two reasons. First, the Wisconsin data allowed us to analyze higher levels of authority (i.e. the power to hire, fire, and determine) than just supervisory capacity. From our Table 1 (1979a:239) it is clear that men are three to four times more likely than women to have these higher levels of authority, while men are only twice as likely as women to be supervisors. The process by which women could be excluded from these higher levels of authority (and how it might differ from the process of acquiring lower levels of authority) was of great interest to us. Second, the Wisconsin data has very good information on the work experiences of women. Since one of the key issues was the extent to which women's lesser attachment to the labor force and a firm were responsible for their low level of authority in the workplace, it was critical to use good measures of work experience and tenure.

Before moving to a comparison of our models estimated with the Wisconsin sample and the QES sample, it is necessary to discuss the measurement of certain variables in both samples. Bridges and Miller assert "we believe that where the correspondence is inexact, the validity of the QES measures is, on balance, equal, if not superior, to those in the Wisconsin Survey." But the Wisconsin sample contains variables that more accurately measure key concepts such as education, work experience, and tenure. Our education measure is the actual number of years of school completed, whereas theirs is an eight point ordinal scale recoded to the midpoint of the interval. Similarly, while they use information from an eight point ordinal scale recoded to the midpoint of the interval to measure tenure, we have the actual number of months with the current employer. Further, we have measured how many months a person has worked since age 16 through a detailed life history, while they must rely on a response to a single question. In short, we believe that they have overestimated work experience and tenure in the QES data set. To us, our measures of these concepts seem superior to theirs, but we ask the reader to consider these differences and make their

* Direct all correspondence to Neil Fligstein, NORC, 6030 S. Ellis, Chicago, IL 60637.

own judgments. While we find this discussion to be a bit trivial, it is important to raise these issues as it is with these variables (which we believe are measured with different amounts of validity) that the difference in our parameter estimates arise.

The practical implications of this difference in variables can be seen in Bridges and Miller's Table 2. While we show statistically significant positive effects for women but not men on the experience and tenure measures, they show significant effects for men on tenure while the rest of the coefficients are statistically insignificant. Since these two coefficients are the most different between their equation and our equation, it is important to note whose measures are closer to true work experience and tenure. We, in short, believe our measures are more valid. Outside of these two coefficients, even Bridges and Miller are forced to admit their results are similar to ours.

Bridges and Miller state that the major differences are that they find experience and tenure effects are larger for men than for women. However, in their model (Table 2), experience is not statistically different from zero for either sex, and we therefore find it surprising that they say it is larger for men than for women. In addition we believe there is a much simpler explanation than theirs for why they find an effect of children for women and we do not. While we originally argued that our lack of an effect of children on authority levels for women (and a positive effect for men) was an anomaly, recent research has indicated that this anomaly is a statistical regularity. Hill (1980), in a study of the sex differences in the acquisition of supervisory responsibility, using the Panel Study of Income Dynamics (and measures of critical variables which may be superior to those in the Wisconsin or QES sample), find that children do not negatively affect the acquisition of supervisory authority for women, but do have a positive effect for men. Rosenfeld (1980) finds this same pattern using status and wages as the dependent variable. It is our belief that when experience is poorly measured, the children variable often becomes important, but that when experience is measured quite accurately, the children variable becomes unimportant.

In terms of the sample selection problem, Bridges and Miller expand their sample to include all persons regardless of age or education level in Table 3. Their Table 3 produces results similar to ours (excluding the differences that arise from the differential measurement of experience and tenure). Bridges and Miller (1981) begin their discussion of Table 3 by noting the fact that "Education continues to have a larger effect for women than for men after the occupation variables are added to the model, although it is now insignificant for both sexes." We find this statement puzzling. In the bottom half of their Table 3, they have failed to star as significant the male coefficients for education after high school. This error invalidates this conclusion. It should also be noted that they failed to star as significant the coefficient for men of the effect of education on supervisory power in our earlier analyses. If they had noted the statistical significance of these parameter estimates, they would have concluded that there were marked similarities between our models. Further, when they do the decomposition using dummy coding of the model in Table 3 (see the upper right half of Table 5), the decomposition looks remarkably similar to the decompositions we present. The fact that their decomposition using the expanded sample and ours are nearly identical suggests that our basic results replicate using a national data set.

The second criticism leveled at us is that our model is misspecified. From a philosophic point of view, all models are misspecified. The social world is a complex place and there are multiple methods of measurement. In our piece, we chose conventional measures in order to minimize the complexity of the model. Bridges and Miller claim that by splining status into white-collar and blue-collar occupations, they have made an important theoretical breakthrough. The bottom line, however, is that their improvements in the specification of our model do not significantly add to the R^2. While this comparison may at first appear to be difficult to make, one can prove that the models are nested. For both sexes, the increment in R^2 by the expansion of the model is not statistically significant ($F_{2,818} = 2.3$ for males and $F_{2,414} = 1.2$ for females). Since they do not significantly add to the R^2 and since they offer little substantive rationale for why their models are inherently better specified than ours, we would argue for the more parsimonious model (ours).

Bridges and Miller's final point is that our regression decomposition is not invariant with respect to the dummy variable coding scheme used to quantify the effects of the categorical independent variables. But they fail to see that substantive interpretability and numerical invariance are not always one and the same. So, having demonstrated that the rates component of the decomposition is invariant under effect coding of the categorical variables, Bridges and Miller consider the problem resolved. However, this is not the case, and a closer examination of their decomposition reveals that numerical invariance is obtained at the expense of substantive interpretability. In short, their solution is less attractive than ours.

To show this, we begin with the observation

that the compositional component is not altered by the choice of a coding scheme (this result is proved in the Appendix). Thus, the difference between the two methods, dummy coding and effect coding, hinges only on the manner in which terms are allocated between the rates and the interactions.

In our decomposition, the rates component may be defined as either

$$[(b_{2m} - b_{1m}) - (b_{2f} - b_{1f})] \times [N_{2f}/(N_{1f} + N_{2f})] + (c_m - c_f)\,\bar{Z}_f \qquad (1)$$

or

$$[(b_{1m} - b_{2m}) - (b_{1f} - b_{2f})] \times [N_{1f}/(N_{1f} + N_{2f})] + (c_m - c_f)\,\bar{Z}_f \qquad (2)$$

where the notation is that employed by Bridges and Miller. To interpret the decomposition, observe that the only difference between Equations 1 and 2 is in the first term. Also, for the sake of concreteness, suppose that the m subscript references males, the f subscript references females, and y is a measure of authority, and x (the categorical variable) is marital status. If marital status is coded 0 at level 1 (absence) and 1 at 2 (presence), one obtains the component in Equation 1. The component itself may then be interpreted as the average amount of additional authority women would attain if (a) they received the male return to z, while maintaining their own characteristic level of z, and if, in addition, (b) women received the same relative return to marriage that men receive, and $[N_{2f}/(N_{1f} + N_{2f})] \times 100\%$ of the women are married. On the other hand, in Equation 2 women receive the same relative returns as unmarried men, and $[N_{1f}/(N_{1f} + N_{2f})] \times 100\%$ of the women are unmarried. In either case, our component admits a clear and meaningful interpretation, despite the fact that its value depends upon the coding of the dummy variable.

Bridges and Miller's rates, however, do not admit this clear interpretation. As demonstrated in the Appendix, their component may be reexpressed as

$$[(b_{2m} - b_{1m}) - (b_{2f} - b_{1f})] \times [N_{2f}/(N_{1f} + N_{2f})] + (c_m - c_f)\,\bar{Z}_f - (1/2)[(b_{2m} - b_{1m}) - (b_{2f} - b_{1f})], \qquad (3)$$

i.e., their component is our Equation 1 with the third term added. Hence, if effect coding is used, one half of the difference between men and women in the relative returns to marriage is arbitrarily subtracted from the rates component (in Equation 1) and added to the interaction component. Thus, if the additional term is positive, the rates component is reduced and the interaction component is increased. This is exactly what occurs in Bridges and Miller's Table 5 on the right side. Our rates component is conditional on the choice of an omitted group, and once the choice has been made, the rate may be interpreted in a traditional manner. It should be clear that by arbitrarily adding the third term, one loses this interpretation. While Bridges and Miller may have gained numerical invariance, they have done so at the expense of substantive interpretability. Since the decompositions in Table 5A are more interpretable than the decompositions in Table 5B, and hence more desirable, Bridges and Miller's approach to the problem is misinformed.

CONCLUSION

We feel that the issues raised by Bridges and Miller have neither undermined nor extended our earlier work. While we are glad to see that our results are sustained using other data and we are interested in useful respecifications of the functional form of variables, we hope that future research will concentrate on redefining how one measures position in the workplace and on the processes, both individual and institutional, that determine one's position in the workplace.

APPENDIX

In this appendix, we establish (a) the invariance of the compositional component under the choice of coding scheme and (b) the relationship between our rates components and the component proposed by Bridges and Miller, as given by Equation 3 in the text. For the sake of continuity, Bridges and Miller's framework and notation is utilized throughout. It should be noted, however, that the results are easily generalized beyond the two-variable case considered herein.

To establish the invariance of the compositional component, we begin by showing invariance under dummy variable coding. If level 1 is set to 0, the compositional component may be expressed as

$$[(N_{2m}/(N_{1m} + N_{2m})) - (N_{2f}/(N_{1f} + N_{2f}))] \times [b_{2f} - b_{1f}] + (\bar{Z}_m - \bar{Z}_f)c_f. \qquad (A.\ 1)$$

If level 2 is set to 0, the appropriate expression is

$$[(N_{1m}/(N_{1m} + N_{2m})) - (N_{1f}/(N_{1f} + N_{2f}))] \times [b_{1f} - b_{2f}] + (\bar{Z}_m - \bar{Z}_f)c_f. \qquad (A.\ 2)$$

However, since

$$[(N_{1m}/(N_{1m} + N_{2m}) - (N_{1f}/(N_{1f} + N_{2f}))] = -1[(N_{2m}/(N_{1m} + N_{2m})) - (N_{2f}/(N_{1f} + N_{2f}))] = w,$$

Equation A. 2 reduces to A.1.

Under effect coding, with level 1 = −1, the composition component is given by

$$[(N_{2m} - N_{1m})/(N_{1m} + N_{2m}) - (N_{2f} - N_{1f})/(N_{1f} + N_{2f})][b_{2f} - b_{.f}] + (\bar{Z}_m - \bar{Z}_f)c_f\,. \qquad (A.\ 3)$$

If level 2 = −1, the appropriate expression is given by

$$[(N_{1m} - N_{2m})/(N_{1m} + N_{2m}) - (N_{1f} - N_{2f})/(N_{1f} + N_{2f})][b_{1f} - b_{.f}] + (\bar{Z}_m - \bar{Z}_f)c_f\,. \qquad (A.\ 4)$$

Since

$$b_{.f} = (b_{1f} + b_{2f})/2,\ b_{2f} - b_{.f} = (b_{2f} - b_{1f})/2,$$
$$\text{and } b_{1f} - b_{.f} = (b_{1f} - b_{2f})/2,$$

so that Equation A.3 may be expressed as

$$[(N_{2m}/(N_{1m} + N_{2m})) - (N_{2f}/(N_{1f} + N_{2f}])) \times [(b_{2f} - b_{1f})/2] + [(N_{1f}/(N_{1f} + N_{2f}) - (N_{1m}/(N_{1m} + N_{2m}))][(b_{2f} - b_{1f})/2] + (\overline{Z}_m - \overline{Z}_f)c_f. \quad (A.5)$$

Next, observe that

$$[(N_{1f}/(N_{1f} + N_{2f})) - (N_{1m}(N_{1m} + N_{2m}))] = -w;$$

The equivalence of Equations A.5 and A.1, and hence of A.3 and A.1, is now immediate. We now have only to show that A.3 and A.4 are equivalent. To see this, we reexpress A.4 as

$$[(N_{1m}/(N_{1m} + N_{2m})) - (N_{1f}/(N_{1f} + N_{2f}))] \times [(b_{1f} - b_{2f})/2] + [(N_{2f}/(N_{1f} + N_{2f})) - (N_{2m}/(N_{1m} + N_{2m}))][(b_{1f} - b_{2f})/2] + (\overline{Z}_m - \overline{Z}_f)c_f. \quad (A.6)$$

The expression A.6, then, is equal to

$$2w[(b_{1f} - b_{2f})/2] + (\overline{Z}_m - \overline{Z}_f)c_f = -w(b_{2f} - b_{1f}) + (\overline{Z}_m - \overline{Z}_f)c_f. \quad (A.7)$$

From this expression, the equivalence of A.7 and A.5, and hence A.3 and A.4, is obvious. Thus, the compositional component is invariant under the alternative coding schemes.

To establish the relationship between our rates and the effect coded rates, as given by Equation 3, we begin with the effect coded rate component

$$[(b_{2m} - b_{.m}) - (b_{2f} - b_{.f})][(N_{2f} - N_{1f})/(N_{1f} + N_{2f})] + (c_m - c_f)\overline{Z}_f. \quad (A.8)$$

Reexpressing $b_{.m}$ and $b_{.f}$ in terms of b_{1m}, b_{2m}, b_{1f} and b_{2f}, and substituting into Equation A.8 yields the expression

$$(1/2)[(b_{2m} - b_{1m}) - (b_{2f} - b_{1f})][N_{2f}/(N_{1f} + N_{2f})] - (1/2)[(b_{2m} - b_{1m}) - (b_{2f} - b_{1f})] \times [1 - (N_{2f}/(N_{1f} + N_{2f}))] + (c_m - c_f)\overline{Z}_f,$$

from which it is clear that Equation A.8 is equivalent to

$$[(b_{2m} - b_{1m}) - (b_{2f} - b_{1f})][N_{2f}/(N_{1f} + N_{2f})] + (c_m - c_f)\overline{Z}_f - (1/2)[(b_{2m} - b_{1m}) - (b_{2f} - b_{1f})]. \quad (A.9)$$

Equation A.9 is identical to Equation 3 of the text.

REFERENCES

Hill, M.
1980 "Authority at work: how men and women differ" in G. T. Duncan and J. M. Morgan (eds.), Five Thousand American Families—Patterns of Economic Progress Vol. VIII. Ann Arbor, Michigan: University of Michigan Press.

Rosenfeld, Rachel
1980 "Race and sex differences in career dynamics." American Sociological Review 45 (August); 583–609.

Wolf, Wendy C. and Neil D. Fligstein
1979a "Sex and authority in the workplace: the causes of sexual inequality." American Sociological Review 44 (April); 235–252.
1979b "Sexual stratification differences in power in the workplace." Social Forces 58 (Sept.); 94–107.

SOCIAL CLASS AND EMOTIONAL DISTRESS*

(COMMENT ON KESSLER AND CLEARY, *ASR*, JUNE 1980)

BETH E. VANFOSSEN

JOHN J. SPITZER

State University of New York at Brockport

DOLORES J. JONES

Western Michigan University

Kessler and Cleary address the important question of why lower-status people have higher rates of psychological distress. Following a comprehensive review of competing theories, the authors reject the hypothesis that greater exposure to stressful life experiences is the major cause. They argue, instead, that people in the lower classes are more distressed because they are more responsive to stress: that is, more likely to develop symptoms of distress when exposed to a problematic life experience. People in the lower classes and people who have not been socially mobile, according to Kessler and Cleary, lack coping resources—"social competence"—which, in turn, makes them emotionally more vulnerable to unpleasant life events. While analytically creative, the paper contains several problems which make the conclusions not entirely convincing.

(1) The indeterminacy of results when controlling for level of exposure to stress. Kessler and Cleary recognize that other researchers have been pursuing the hypothesis that greater exposure to stressful life events accounts for higher rates of distress in the lower classes. They hypothesize, however, that people in lower-class positions are more likely than others to exhibit symptoms of distress at any given level of exposure. Their data come from the New Haven longitudinal survey of mental disorder (Myers et al., 1972, 1974). Myers et al. (1974) repeatedly indicate (pp. 195, 196, 198, 199, 202) that when undesirable life events are statistically controlled, the association between social class and emotional distress disappears, regardless of the type of statistical test used. Myers et al. conclude (1974:202), "the greater amount of psychiatric distress found in the lower class in community studies is due to the uneven distribution of life events measured by a scale of Desirability-Change. . . . the data demonstrate the importance of social and impersonal forces frequently external to the individual in influencing

* Direct all communications to Beth E. Vanfossen, Department of Sociology, State University of New York, Brockport, New York 14420.

psychological status." Kessler and Cleary do not explain in their paper why they reject the earlier conclusion based on the same data set, nor do they offer evidence that the relationship between class and distress exists when life events are controlled. In an earlier paper (Kessler, 1979:264–5), the Myers et al. finding is reported, but dismissed with the statement that the Myers et al. analysis failed to consider impact differentials.

(2) The contamination of the independent, control, and dependent variables. An important segment of the Kessler and Cleary paper reports the results of regression analyses relating emotional distress scores (Gurin2) to three independent variables: the level of emotional distress that had existed two years earlier (Gurin1), the presence of physical health problems (HP), and the number of undesirable life events that had occurred in the prior two years (WTUE). Separate analyses were performed for the lower class, the working class, and the middle and upper classes combined. The dependent variable, emotional distress at Time 2 (Gurin2), was measured by the Gurin scale, which in turn was based on a symptom scale developed by Macmillan (1957) and modified by Gurin et al. (1960). It consists of 20 items, a number of which appear to be measures of ill health rather than anxiety or depression as the authors assume.[1] While Kessler and Cleary took great care to avoid mixing psychosomatic symptoms with health symptoms in their measure of "physical health problems," their measure of emotional distress contains six items that an unhealthy person would clearly be more likely to check: the presence of a health problem, feeling healthy enough to do things, being bothered by pains and ailments, finding that ill health interferes with work, being bothered by the heart beating hard, and being bothered by shortness of breath when not working hard. The inclusion in Gurin2 of these health items is consequential because one of the independent variables—the physical health problems index—is a score based on weighted responses for the presence of 16 health problems such as cancer, operations, injury, and diabetes. Kessler and Cleary report, as we might expect, that HP is significantly related to Gurin2, particularly for skilled and unskilled workers and/or their spouses.

Because Gurin2 is heavily based on the reporting of signs of ill health, it measures conditions more characteristic of the poor than of the affluent. Lower-class status brings with it a higher incidence of ill health and of chronic illness (U.S. National Center for Health Statistics, 1965). The incidence of diabetes, heart disease, and cancer, in particular, is higher in the lower social classes (Ornati, 1966). Thus, given the heavy reliance of the Gurin scale on items relating to health problems, Kessler and Cleary's report that respondents in the lower social classes have higher distress scores on Gurin2 is not surprising.

Another possibility of contamination lies in a class bias inherent in the measurement of the independent variable, "Undesirable Life Events" (WTUE). While some of the events that make up this scale are probably unrelated to class, others measure events that more frequently affect people in the lower classes: suffering serious injury or accident or being out of work for over a month, fired, foreclosed, detained in jail, or laid off temporarily. The contamination results from the fact that both the Gurin scale and the WTUE are measuring objective conditions that are class related. If both scales include a number of items more likely to be checked by people in the lower classes, then there should be, as reported, a greater association between the two scales for the lower classes than for the higher classes. The scores obtained by the higher-class respondents on the two independent variables would be based more heavily on the remaining less-class-biased items. It is thus unclear what is really being indicated by Kessler and Cleary's regression results. The question is raised of what would be the effect on the class-distress relationship of removing from Gurin2 all items dealing with physical health symptoms, or from the WTUE scores those events experienced mainly by the poor and unskilled.

(3) The cumulative effect of multiple disasters. Previous research has indicated that the more rigorous and severe the external situation, the less significant are individual characteristics in determining the likelihood and nature of response (Rabkin and Struening, 1976:1018). Under conditions of multiple disasters, additional hardships become progressively more stressful. Kessler and Cleary report that the WTUE scores for lower classes are four times higher than those of the top classes (Table 3). In Footnote 9, they address the concern expressed by one reviewer that the higher "distress" scores of the lower classes might mainly reflect the cumulative effect of multiple disasters, rather than a greater emotional responsiveness to hardship. They perform a series of polynomial regressions to see if the association between WTUE and Gurin2 increases at any point on the regression curve.

[1] That there are different conceptual dimensions being measured in the Gurin scale was shown by Depner and Kulka (1980), whose factor analysis of the twenty items showed the presence of two factors which they termed "physical ill health" and "psychological anxiety."

However, the equation adopted to test the curvilinearity of the relationship seems inappropriate. First, the equation lacks an intercept term; the resulting homogeneous equation may conceal a curvilinear relationship, whereas an equation with an intercept term allowed may reveal one. Second, one wonders why HP (physical health problems) was omitted from the equation, when the original behavioral equation includes HP. Thirdly, if it is assumed that the three social class groups are different with respect to WTUE, each group should have had separate equations. The aggregation of three "different" groups into the one sample may have obscured a nonlinear relationship. Finally, substituting the mean values[2] of the independent variables into the right hand side of the equation obtains

$$\begin{aligned} G_2 &= .496(71.97) - 1.5(84.5) \\ &\quad + 1.021(7142) \\ &= 7201 \end{aligned}$$

Since G_2 averages around 72, how can this estimated equation be correct?

(4) The omission of other significant variables. The higher WTUE beta coefficient in the regression performed with the lower-class data (Table 3) may be a reflection more of the fact that Class V people have fewer resources to deal with difficulties than of their being less competent copers. Suppose, for example, that the WTUEs experienced by Class V people are associated with other, more debilitating problems, such as financial difficulties. This would be true for those respondents who experienced the unpleasant life events of being fired, unemployed, or foreclosed. We know that Class V people have fewer objective resources (money and status) to deal with the onset of such problems. The loss of a job may have more severe economic consequences for the Class V people than for Class I people, resulting in higher average emotional distress scores for the Class V people, socially competent or not. The same would be true for the occurrence of injury and even the loss of a spouse. Kessler and Cleary's footnote 5 indicates that financial problems are indeed related to Gurin scores. If our reasoning here is correct, then WTUE would produce a lower beta for Class I–III respondents than for Class V respondents because the impact of undesirable life events is diminished by the greater economic resources of the higher classes. The possibility reminds us that if other unmeasured variables are correlated with both the independent and dependent variables (financial distress with WTUE, and also with Gurin2), or if an unmeasured variable is an intervening variable (the impact of events being mediated by availability of financial resources), then the betas in a regression equation are biased in an unknown way, and should not be interpreted as though they represent the direct causal paths.

(5) The possibility of alternative interpretations of the Kessler-Cleary model. The model specified by Kessler and Cleary for estimation

$$\begin{aligned} G_t &= a_0 + AG_{t-1} + a_2HP_t \\ &\quad + a_3WTUE_t \end{aligned} \tag{1}$$

where t refers to the current year and $t-1$ to the previous year, can be interpreted in a radically different fashion. Let us assume that current Gurin scores are influenced not only by current HP and WTUE indices, but also to a lesser extent by the values of these indices in years prior to measurement. This assumption would be reasonable, for example, for someone whose current anxiety level has been heightened by a diagnosis of cancer three years ago. Even if the disease is now in remission and does not enter into a current HP score, it might elevate anxiety levels. Similarly, an operation or loss of a relative several years ago that might still have physical or emotional effects would not be recorded in current WTUE scores or current HP scores. Thus, not only current, but lagged values of HP and WTUE, with influence declining as they become more temporally distant, may explain current Gurin scores.

This type of model, called a distributed-lag model (Koyck, 1954), results in an estimable form identical to Equation 1 (see Appendix). Under this interpretation, the coefficient on G_{t-1} is the weight to be assigned to lagged values of HP and WTUE. The larger A is, the more important are lagged values of HP and WTUE in explaining G_t, and consequently, the less important are the current values of HP and WTUE. From Kessler and Cleary's Table 3, the Koyck interpretation implies that stress in social position I–III is more strongly influenced by the past than for either social position IV or V. Alternately, social positions IV and V are more influenced by immediate stress (current HP and WTUE) than social position I–III. Since the specified model is capable of being interpreted in two radically different ways, Kessler and Cleary's conclusion becomes questionable. Their model is not identified explicitly to exclude the Koyck interpretation.

---

[2] The average value of WTUE-squared is underestimated by squaring the average of WTUE. Note that the variance of WTUE, $= V(WTUE) = \Sigma\,(WTUE)^2/N - (\overline{WTUE})^2$. Thus, the average of $\underline{(WTUE)^2} = \Sigma\,(WTUE)^2/N = V(WTUE) + (\overline{WTUE})^2$. An estimate of $\Sigma\,(WTUE)^2/N$ by $(\overline{WTUE})^2$ underestimates the term by $V(WTUE)$.

Under the Koyck interpretation, the lagged dependent variable, G_1, is not independent of the (autocorrelated) disturbance term, and ordinary least squares estimates of the coefficients are no longer either unbiased or consistent. Thus, even the size and statistical significance of the coefficient estimates in the Kessler-Cleary model are suspect.

(6) Evidence for the influence of "social competence." Kessler and Cleary find that emotional distress is less likely to be produced by undesirable life events in people who have been upwardly socially mobile than in people who have been stable or downwardly mobile. The regression coefficient for this finding (Table 4) is not particularly large, but it is statistically significant. The finding, incidentally, is inconsistent with some earlier research (see Kessin, 1971, for one review). From this finding, they suggest that the experience of success associated with upward mobility creates "social competence," characterized by active problem-solving styles.

To substantiate the idea, Kessler and Cleary make a series of deductions. For instance, they suggest that women who are upwardly mobile through marriage are less likely than those who have been upwardly mobile through success in the business world to be "socially competent." Rather than directly comparing women mobile in business to those mobile in marriage, Kessler and Cleary confound employment status with sex by comparing chief breadwinners (both male and female) to not–chief breadwinners, thus making the findings uninterpretable.

In comparing men and women, Kessler and Cleary find that mobile, male breadwinners are less likely to check off the symptoms on the Gurin scale following the occurrence of undesirable life events than are mobile, female nonbreadwinners. They state (p. 473), "these various results suggest that the comparatively lower reaction to stressful life experiences of upper status people in this sample is due to the presence among them of extremely competent copers who have been upwardly mobile from lower class origins."

That men have lower rates of depression and anxiety than women is consistently found in studies of sex differences in emotional distress (Dohrenwend and Dohrenwend, 1976). Why this is so is not clear, and a number of hypotheses have been offered (Weissman and Klerman, 1977; Rosenfield, 1980). The idea that males and chief breadwinners are socially more competent is only one of a number of plausible interpretations. Kessler and Cleary's conclusion would have been more persuasive had they presented indicators of social competence directly, rather than compared the emotional distress levels of males and females who differ on many other things (including the probability of experiencing financial difficulty).

Their conclusion is weakened further by the finding that mobility does not seem to modify the relationship between physical health problems and distress. If the upwardly mobile are superior copers, then one would expect that they would deal more effectively with their health problems than those who are stable or downwardly mobile.

In sum, Kessler and Cleary's conclusion that social class differences in emotional distress are determined by class differences in responsiveness to life problems is not convincing. They have failed to take into account the disappearance of the class-distress relationship when life events are controlled and have failed to provide any direct measure of social competence. Moreover, both dependent and independent variables may be spuriously contaminated, and important concomitant variables may have been omitted. Lastly, because their model allows an alternative interpretation that leads to conflicting conclusions, its validity must be questioned.

APPENDIX

Let the original model be specified by

$$\begin{aligned} G_t = {} & B_0 + B_1(HP_t + AHP_{t-1} \\ & + A^2HP_{t-1} + \ldots \\ & + A^THP_{t-T}) + B_2(WTUE_t \\ & + AWTUE_{t-1} + A^2WTUE_{t-2} \\ & + \ldots + A^TWTUE_{t-T}) \end{aligned} \tag{2}$$

where $0 < A < 1$. Powers of A are a geometrically declining series. The interpretation of Equation 2 is that HP and WTUE both influence G_t, but that the influence of scores on G_t declines, or decreases with time.

Equation 2 does not lend itself to easy estimation; too many observations are lost by the multiple lags, and the problem of multicollinearity becomes severe. Thus, a reformulation of Equation 2 is called for. We write Equation 2 lagged one period, and multiply this result by A, obtaining

$$\begin{aligned} AG_{t-1} = {} & AB_0 + B_1(AHP_{t-1} \\ & + A^2HP_{t-2} + \ldots \\ & + A^{T+1}HP_{t-T-1}) \\ & + B_2(AWTUE_{t-1} + A^2WTUE_{t-2} \\ & + \ldots + A^{T+1}WTUE_{t-T-1}) \end{aligned} \tag{3}$$

Equation 3 only asserts that if 2 holds for the current period, the same relationship should also hold for the previous period. We now subtract Equation 3 from Equation 2 to obtain

$$\begin{aligned} G_t = {} & (B_0 - B_0A) + AG_{t-1} \\ & + B_1HP_t + B_2WTUE_t \\ & - B_1A^{T+1}HP_{t-T-1} \\ & - B_2A^{T+1}WTUE_{t-T-1} \end{aligned} \tag{4}$$

Since $0 < A < 1$, large powers of A, such as A^{T-1}, are close to zero, and the last two terms are dropped, resulting in Equation 1 with $a_0 = B_0 - B_0A$, $B_1 = a_2$, and $B_2 = a_3$. The equation to be estimated, resulting from the Koyck transformation, is identical to Equation 1, but its interpretation is not that of Kessler and Cleary.

REFERENCES

Depner, Charlene E. and Richard Kulka
1980 "Sex differences in psychological distress and use of professional help as a function of social and social-psychological factors." Project No. 462213. Ann Arbor: Institute for Social Research.

Dohrenwend, Bruce P. and Barbara S. Dohrenwend
1976 "Sex differences in psychiatric disorders." American Journal of Sociology 81:1147–54.

Gurin, Gerald, Joseph Veroff, and Sheila Feld
1960 Americans View Their Mental Health. New York: Basic.

Kessin, Kenneth
1971 "Social and psychological consequences of intergenerational occupational mobility." American Journal of Sociology 77:1–18.

Kessler, Ronald C.
1979 "Stress, social status, and psychological distress." Journal of Health and Social Behavior 20:259–72.

Kessler, Ronald C. and Paul D. Cleary
1980 "Social class and psychological distress." American Sociological Review 45:463–78.

Koyck, Leendert M.
1954 Distributed Lags and Investment Analysis. Amsterdam: North Holland.

Macmillan, Allister M.
1957 "The health opinion survey: technique for estimating prevalence of psychoneurotic and related types of disorder in communities." Psychological Reports 3:325–9.

Myers, Jerome K., Jacob J. Lindenthal, and Max P. Pepper
1974 "Social class, life events and psychiatric symptoms: a longitudinal study." Pp. 191–206 in B. P. and B. S. Dohrenwend (eds.), Stressful Life Events. New York: Wiley.

Myers, Jerome K., Jacob J. Lindenthal, Max P. Pepper and David R. Ostrander
1972 "Life events and mental status: a longitudinal study." Journal of Health and Social Behavior 13:398–406.

Ornati, Oscar
1966 Poverty Amid Affluence. New York: Twentieth Century Fund.

Rabkin, Judith G. and Elmer L. Struening
1976 "Life events, stress, and illness." Science 194:1013–20.

Rosenfield, Sarah
1980 "Sex differences in depression: do women always have higher rates?" Journal of Health and Social Behavior 21:33–42.

U.S. National Center for Health Statistics
1965 "Chronic conditions and activity limitation, United States, July 1961–June 1963." Ser. 10, no. 17. Washington, D.C.: U.S. Government Printing Office.

Weissman, Myrna M. and Gerald L. Klerman
1977 "Sex differences and the epidemiology of depression." Archives of General Psychiatry 34:98–111.

REPLY TO VANFOSSEN ET AL.*

RONALD C. KESSLER
University of Michigan

PAUL D. CLEARY
Rutgers University

We think debate about class differences in responsiveness to stress can be stimulating and productive, but we take issue with Vanfossen et al.'s characterization of our theoretical arguments and disagree with their critique of our analytic methods. Because the major part of their comment is organized around six methodological issues, we will address those questions first and then respond to the substantive criticism of our work that appears to underlie most of the other comments.

(1) Our purpose in writing the paper was to describe an analytic strategy for disentangling contending interpretations of the fact that undesirable life events have their most damaging psychological effects on lower class people. The existence of this interaction between life events and class in predicting psychological distress was first documented nearly two decades ago in the Midtown Manhattan study (Langner and Michael, 1963) and has been replicated several times since then (Dohrenwend, 1973; Brown and Harris, 1978; Markush and Favero, 1974).

In the face of this consistent evidence that an interaction exists, we are perplexed by the first criticism raised by Vanfossen and her associates: that we ignored the fact that Myers et al. (1974), working with the same data we analyzed, showed that the relationship between class and distress disappears when the additive influence of life events is controlled. Our critics apparently fail to appreciate that the Myers et al. model is misspecified because it does not include the interaction between WTUE and class that we know exists in these data.

(2) The next point raised by Vanfossen et al. is that the differential effects of HP and WTUE on Gurin scores in the different social classes represent "contaminations" of the independent and dependent variables.

* Direct all correspondence to Ronald C. Kessler, Department of Sociology, University of Michigan, 3012 LSA Building, Ann Arbor, MI 48109.

Work on this response was supported by NIMH grant R01MH34479. We appreciate the helpful comments and suggestions of David Mechanic, James McRae, and Gillian Stevens.

We begin by considering the HP-Gurin relationship. As Vanfossen et al. note, the Gurin scale contains a series of items measuring ill health. These were included in the scale because they are among the symptoms most frequently reported by psychiatric patients in treatment, reflecting the fact that they are not only indicative of physical health problems but also of psychophysiological distress. In our paper we noted that the inclusion of these symptoms created a potential problem for interpreting the effects of HP on Gurin and were careful to include in the HP index only serious types of physical health problems that are clearly not psychosomatic in origin.

The criticism is that the inclusion of these items in the Gurin scale "is consequential" because health problems are "more characteristic of the poor than the affluent." But this is irrelevant. The fact that poor people are more likely to be in poor health than affluent people does not in any way explain why the occurrence of a physical health problem (HP) is associated with a higher increase in subjectively experienced symptoms of ill health among lower/working class respondents than among middle/upper class ones.

Vanfossen et al. make the same error when they assert that the WTUE-Gurin relationship is "contaminated." Here the argument is that "both the Gurin scale and the WTUE are measuring objective conditions that are class-related. If both scales include a number of items more likely to be checked by people in the lower classes, then there should be, as reported, a greater association between the two scales for the lower classes than for the higher classes." This conclusion is incorrect. To make a trivial, but possibly illuminating, parallel, we know that height and weight are both associated with gender. But this does not mean that the height-weight relationship is stronger for men than women. Indeed, nothing whatever about the relative magnitudes of this relationship among men and women separately can be inferred from knowledge of the respective means. Similarly there is no reason to think that the relationship between WTUE and Gurin will be stronger in the lower classes because class is negatively associated with WTUE and Gurin. The "contamination" our critics allude to does not exist.

(3) Vanfossen and associates correctly note that the intercept is missing in the regression equation reported in footnote 9 of our paper. It was inadvertently omitted from the published version of the paper. They also criticize us for not controlling HP when we estimated nonlinearities in the WTUE-Gurin relationship. We have no reason to think that variations in HP would mask a nonlinearity in the relationship of WTUE to Gurin. Nor do our critics offer any argument of this sort. Therefore, we fail to see the point of this criticism.[1]

They also criticize us for failing to estimate the nonlinearity within the separate social classes, since we "assumed that the three social class groups are different with respect to WTUE." This criticism is ill-conceived. The very reason we investigated the possibility of a nonlinear association was that we wanted to assure ourselves that the estimated interaction was not due to a nonlinearity in the combined data. Only after we found that no meaningful nonlinearity of this sort existed did we assume that the interaction was real. Estimating the nonlinearity in separate classes would serve no purpose.

(4) As mentioned above, our primary intent in writing this paper was to present an analysis strategy to disentangle contending interpretations of the class-stress interaction. In their fourth criticism, Vanfossen et al. present us with just such a contending set of interpretations, but miss the critical virtue of our approach in helping decide between them.

Their argument is that the differential impact of WTUE may reflect the fact that Class V people experience more difficulties rather than that they are less competent copers. They then go on to argue that a variety of resource deficits that are more characteristic of the poor could exacerbate the emotional effects of life events.

Vanfossen et al. apparently are unaware that our analysis approach is designed explicitly to test this type of possibility. We very clearly discussed this as part of the "social resources" perspective on responsiveness.

---

[1] A mistake was made in tabulating the WTUE coefficient in the footnote, as can be seen by noting that the ratio of metric to standardized coefficients in the footnote differs from the ratio in Table 4. The correct equation is reported below. In compliance with Vanfossen et al.'s suggestion, we have controlled for HP as well as for age, sex, and marital status.

$$
\begin{aligned}
G_2 = {} & .440{+}2 + .423\ (.452)\ G_1 \\
& - .693{-}2\ (-.217)\ \text{WTUE} \\
& + .194{-}5\ (.092)\ \text{WTUE}^2 \\
& - 1.152\ (-.175)\ \text{HP} \\
& + \text{controls};\ R^2 = .324.
\end{aligned}
$$

Coefficients outside the parentheses are expressed in scientific notation to take account of the widely differing metrics. Coefficients inside the parentheses are standardized. The WTUE^2 coefficient is significant at $p = .104$. All other coefficients are significant at $p = .000$. As we see, then, the effect of controlling for HP is to inflate somewhat the positive effect of WTUE^2, which argues even more strongly for the interpretation presented originally in the footnote.

> The social resources perspective emphasizes differences in access to resources that can be effective in dealing with stress, thus suggesting that one's current class position . . . will be associated with low resistance to stress. Consequently, *if we find that an observed class difference in responsiveness is not, at least in part, traceable to differences in class destinations, controlling for origins and mobility experiences, a social resources explanation is discredited.* [Kessler and Cleary, 1980:467, emphasis added]

As reported in Table 4 of our paper, the interaction of WTUE with class destinations was insignificant. We concluded that "these findings are inconsistent with . . . the social resources perspective. Instead, we find that upwardly mobile people are comparatively unresponsive to stress." Since upwardly mobile people are most common in the I–III subsample, higher responsiveness is associated with lower status. We went on to state that this pattern is consistent with more than one substantive interpretation. However, it is completely inconsistent with the interpretation advanced by Vanfossen, Spitzer, and Jones. On the basis of this evidence, which our critics neither dispute nor acknowledge, we find their alternative interpretation uncompelling.

One point raised in this part of their critique is important to address. They mention that "if an unmeasured variable is an intervening variable (the impact of events being mediated by availability of financial resources), then the betas in a regression equation are biased in an unknown way, and should not be interpreted as though they represent the direct causal paths." The measures of social origins, destinations, and mobility experiences that we used to explain responsiveness to stress are proxies for the class-related resources and resource deficits that are responsible for class differences in responsiveness.

As stated in our paper, this proxy approach is much less satisfactory than an approach that explicitly measures the resources and deficits thought to be important. But we know too little about class differences to carry out a more explicit analysis. Our approach provides a beginning—nothing more—by helping us trim down the possibilities to investigate in more concrete detail. As stated in our paper, "the strength of this decomposition approach lies in its ability to reject particular interpretations. We learn much less here about the forces that are at work than about those that are not. . . . Furthermore . . . [the decomposition approach] can be augmented (in certain instances) with other data so that we can arrive at some relatively clear understanding of the causal processes at work in a particular instance." Thus we think it is unreasonable for Vanfossen et al. to assert that our regression coefficients are "interpreted as though they represent the direct causal paths."

(5) Vanfossen et al. next suggest that a distributed lag formulation might more accurately capture the causal influences of HP and WTUE on Gurin scores than the single lag specification we estimated. They note correctly that our model "is not identified explicitly to exclude" this different specification. And then, after making some remarks intended to show why this alternative specification is plausible, they assert that because of our failure to estimate their specification our estimates "are no longer either unbiased or consistent" and that "even the size and statistical significance of the coefficient estimates in the Kessler-Cleary model are suspect."

The criticism boils down to the trivial claim—one that can be made against any model—that the results are wrong if the specification is wrong. The distributed lag specification they propose is not, as they imply, merely a different interpretation of the same model specified in our analysis. Rather, it is a model that resembles ours in some ways, but differs radically in the assumptions made about the presence of serial correlation in the specification errors. Furthermore, the model they derive is only one of a broad class of distributed lag models, one that is particularly implausible on substantive grounds.[2]

We will not discuss distributed lag specifications because there are substantive reasons why the single lag specification is more appropriate. Several case-control studies of the onset of mental health problems have investigated the issue of appropriate time lag in the relationship between life events and emotional distress (Brown and Harris, 1978; Paykel, 1978; Surtes and Ingham, 1980), and have shown that the effects of events more than a year in the past are negligible. This is not to say that a person might not have a breakdown some years after the occurrence of a traumatic event due to a relatively minor precipating incident that recalls that event. However, in terms of the overall association of life events with emotional difficulties this type of association is relatively unimportant. Furthermore, since exposure to traumatic life events of this sort is not stable over time, omitting these delayed effects from one's model will not bias the estimated effects of more proximate life

[2] There are a variety of other distributed lag models that are more plausible than this one and which can be evaluated in comparison to our model by means of sensitivity analysis. Discussion of these models can be found in Kessler and Greenberg (1981).

events. This is also true of the distributed lag model proposed by Vanfossen et al., since there is no lagged endogenous variable in their prediction equation. They are consequently incorrect when they say that coefficients are biased when OLS is used to estimate this model.

Even on an intuitive level the model they propose is implausible. It assumes that an event several years in the past can still influence change in distress but that the level of distress at the beginning of the panel has no effect on the distress reported at the end of the panel. This means, in the context of the New Haven panel, that the level of distress a respondent exhibits in 1967 is assumed not to influence his level of distress in 1969, while an event that occurred in, say, 1965 can influence 1969 distress. The model we estimated, in contrast, assumes that 1967 distress can influence 1969 distress, that life events experienced over the past year can influence 1969 distress, and that earlier events have negligible effects. These assumptions are, to our minds, intuitively much more plausible than those advanced by Vanfossen and her associates.

(6) The last criticism accepts, conditionally, our finding that the emotional distress produced by undesirable events is less pronounced among the upwardly mobile, but disputes our interpretation.[3] We interpreted the interaction as being consistent with variants of both social-selection and adult-socialization hypotheses, and went on to say that we favored the adult-socialization perspective, because of its consistency "with the literature on human development, which emphasizes the importance of adversity for the creation of competence, especially when this adversity is encountered early in one's life" (Kessler and Cleary, 1972:472).

We stated that our data did not allow us to disentangle the various complex influences that make up this interaction, but we presented two pieces of inferentially supportive evidence. First, it was shown that at least one coping response about which we had information (the use of tranquilizers during times of stress) is less characteristic of upwardly mobile people. Second, we showed that the interaction of WTUE with upward mobility is more pronounced among men than women and among chief breadwinners than people who are not chief breadwinners.

We did not present a breakdown by the cross-classification of sex and employment because the subsamples defined by this cross-classification were extremely small. But we did state that a more detailed specification of this sort would have been preferable. In particular, we stated that "we would expect women who made their ways to the top through their own successes in the business world to be equally as competent (and perhaps more so) as men who had the same successes."

Nowhere in our presentation of these additional specifications did we state, nor even imply, that these small additions to our overall results were anything more than supportive, inferential evidence in favor of the interpretation we offered for the mobility effect. However, in this criticism, Vanfossen, Spitzer, and Jones focused on these peripheral analyses to accuse us of "confounding" employment status with sex and making too much of the consistently documented finding that "men have lower rates of depression and anxiety than women . . ." We object to this criticism on two grounds. First, we made much less of these inferentially supportive results than our critics do. It is true that the results are not as precise as we would have liked. And we agree with Vanfossen and company when they say "Kessler and Cleary's conclusion would have been more persuasive had they presented indicators of social competence directly, rather than compared the emotional distress levels of males and females who differ on many other things (including the probability of experiencing financial difficulty)." However, we think it is unreasonable to fault us for being unable to transcend the limitations of the data we had at our disposal. Second, we did not merely find that men have lower rates of distress than women. We found that the interactive influence of upward social mobility and undesirable events on distress was more pronounced for men than women. While it is admittedly true that this specification could be due to any of several sources, it is not true that all the hypotheses offered to account for sex differences in mental illness are possibilities. This might be a minor point to quibble about, but it exemplifies an annoying tendency on the part of our critics to treat the detailed patterns of association and interpretation we carefully presented in our paper in a casual fashion.

[3] Vanfossen et al. state: "The finding, incidentally, is inconsistent with some earlier research." A paper by Kessin (1971) is cited as an example of conflicting evidence. The Kessin paper reviews the literature on the gross association between social mobility and psychological functioning, which is a much broader topic than the one covered in our paper. We made explicit mention of this larger topic and its irrelevance for our work in footnote 13 of our paper. Apparently our critics overlooked this footnote when they prepared their comments. To our knowledge, no other published study has reported information about differential responsiveness to stress in subgroups defined in terms of mobility experience.

A MORE BASIC CRITICISM

Others have expressed concern that our paper might contribute to a tradition of explaining away lower class miseries by ascribing their cause to personality and value characteristics rather than structural sources. We want to address this deeper concern, for we share it with our critics.

When studying stratification it is easy to slip into a psychological perspective that attributes lack of attainment to personality characteristics such as lack of motivation, lack of competence, etc. This narrow view is particularly dangerous in a study of coping, where discussion of intrapsychic facilities is an inherent part of the endeavor. Therefore, every effort should be made in research of this sort to tie the personality characteristics under investigation to their structural sources. We would like to think that our orientation shows this type of sensitivity. We regret that Vanfossen and her associates did not find it in our paper.

The discussion of social competence in our paper emphasized an adult-socialization perspective. The data revealed, in other words, that heightened responsiveness to the life events included in our inventory was more closely associated with the history of coming to occupy a particular social position than with the current conditions associated with that position. We made no attempt to investigate these conditions in the paper, but not because we felt that our data provided an answer to our question about class differences in mental health. On the contrary, we hoped for no more than to provide a direction for more concrete research about the structural features of mobility associated with this heightened responsiveness to life stress.

We also mentioned in our paper, and it is worth repeating here, that we are not convinced that mobility is equally associated with responsiveness to all types of stressful life experiences. Indeed, a reviewer of our paper encouraged us to focus only on the findings about the WTUE-Gurin relationship and ignore that for the HP-Gurin relationship, but we rejected this suggestion for the very reason that we wanted to demonstrate that the results of our decomposition vary depending on the type of stressor under examination. For WTUE the interaction was associated with upward social mobility; for HP it was associated with social origins. In more recent unpublished work, one of us has documented that some types of events interact most strongly with current social position.[4] Indeed, unpublished analysis of several parallel data sets indicates that the interaction of WTUE with upward mobility found in the New Haven data is not common to most data sets. This work has shown, however, that the overall interaction between a scale of undesirable life events interacts with each of several different indicators of social class (respondent and family income, education, and occupational status) to predict a variety of distress indicators. Work is now proceeding to estimate event-by-event models to determine what events do and do not interact with class to predict distress, and, among events that have more damaging emotional effects on lower class people, for which events these patterns are most closely associated with social origins, destinations, and/or mobility. This type of detailed descriptive question was the one the decomposition strategy was designed to study. Far from attempting to offer the approach as an end point of our study of class differences in responsiveness to stress, we see it as a tool for elaborating in fine detail the patterns of responsiveness that exist across a range of stress experiences. This descriptive elaboration can then serve as a basis for speculation, refinement, and concrete hypothesis testing about the structural determinants of class differences in responsiveness.

It is perfectly clear from the work done so far that a large part of the relationship between social class and emotional distress stems from the greater damage inflicted on the poor by life events. This greater responsiveness is not well understood. But we firmly believe that it is real and that it has to be studied carefully before we are going to understand class differences in mental illness. Our approach does not answer all the questions we have of this pattern. The preliminary analysis presented in the paper describing the strategy did, we hope, raise some substantive questions. We feel that the current work being done elaborating descriptive patterns in more detail will raise other questions and stimulate some thoughts about how to investigate more precise hypotheses about responsiveness.

[4] A draft copy of this unpublished analysis will soon be available from the senior author.

REFERENCES

Brown, George W. and Tirril Harris
1978 Social Origins of Depression. New York: Free Press.

Coles, Robert
1967 Children of Crisis Volume I: A Study of Courage and Fear. Boston: Little, Brown.

Dohrenwend, Barbara S.
1973 "Social status and stressful life events." Journal of Personality and Social Psychology 9:203–14.

Elder, Glen H., Jr.
1974 Children of the Great Depression. Chicago: University of Chicago Press.

Kessler, Ronald C. and P. D. Cleary
1980 "Social class and psychological distress." American Sociological Review 45:463–78.
Langner, Thomas S. and Stanley T. Michael
1963 Life Stress and Mental Health. New York: The Free Press.
Kessler, Ronald C. and David F. Greenberg
1981 Linear Panel Analysis. New York: Academic.
Markush, R. E. and R. V. Favero
1974 "Epidemiologic assessment of stressful life events, depressed mood, and psychophysiological symptoms–a preliminary report." Pp. 171–90 in Bruce P. Dohrenwend and Barbara S. Dohrenwend (eds.) Stressful Life Events. New York: Wiley.
Moriarty, Alice E. and Povl W. Toussieng
1976 Adolescent Coping. New York: Grune and Stratton.
Murphy, Lois Barclay and Alice E. Moriarty
1976 Vulnerability, Coping, and Growth. New Haven: Yale University Press.
Myers, J. K., J. J. Lindenthal, and M. P. Pepper
1974 "Social class, life events and psychiatric symptoms: a longitudinal study." Pp. 191–206 in Bruce P. Dohrenwend and Barbara S. Dohrenwend (eds.) Stressful Life Events. New York: Wiley.
Paykel, E. S.
1978 "Contribution of life events to causation of psychiatric illness." Psychological Medicine 8:245–53.
Pearlin, L. I. and C. Schooler
1978 "The structure of coping." Journal of Health and Social Behavior 19:2–21.
Surtes, P. G. and J. G. Ingham
1980 "Life stress and depressive outcome: application of a dissipation model to life events." Social Psychiatry 15:21–31.

ON THE USE AND ABUSE OF AGGREGATE DATA IN ANALYZING ORGANIZATIONS

(COMMENT ON JAMES R. LINCOLN AND GERALD ZEITZ, *ASR*, JUNE 1980)*

MARSHALL W. MEYER
WILLIAM B. STEVENSON
University of California, Riverside

The Lincoln and Zeitz article is a potentially important contribution to organizational research because it uses methods heretofore overlooked by organizational analysts to address a key problem: that of separating individual effects upon organizational properties from aggregate or structural effects. However, in their haste to resurrect the "earlier conception of bureaucracy" (p. 406), which views growth and elaboration of structure as antithetical to decentralization, Lincoln and Zeitz have overlooked a number of basic theoretical precepts, and their methodology has obscured rather than illuminated the issues involved.

* Direct all correspondence to Marshall W. Meyer, Department of Sociology, University of California, Riverside, CA 92521.

THEORETICAL ISSUES

Lincoln and Zeitz begin and end their article with brief statements about the relationship of decentralization to other organizational properties such as size, division of labor, administrative intensity, professionalization, and uncertainty or change; at these points, the issue of individual versus aggregate properties seems irrelevant. Lincoln and Zeitz could have parsimoniously treated decentralization as a function of other variables describing organizations and estimated the appropriate regression models. That they did not do so indicates that Lincoln and Zeitz were, in fact, testing a different theory, one relating participation in organizations to other properties. The latter theory is most parsimoniously tested at the individual level of analysis, since the dependent variable describes individual persons, as do most of the independent variables in Lincoln and Zeitz's study. Independent variables not easily reduced to the individual level could have been treated as contextual measures. Indeed, it would have been preferable to undertake this analysis at the individual rather than at the aggregate level, since, as is well known, aggregation tends to increase bias in estimates of regression parameters (Hannan et al., 1976).

What is critical but overlooked in the Lincoln and Zeitz article is that decentralization in organizations and participation in decisions are not necessarily the same and may be antithetical. Participation is input to a decision; decentralization, by contrast, is the extent to which persons below the boss actually make decisions. Decentralization thereby removes the boss from decisions. Organizations moving from centralized to decentralized decision making often do so not because too few people are participating in decisions but because too many are. As Williamson (1975) points out, decentralization economizes on bounded rationality. Typically, decentralization, again as measured by the extent to which persons below the boss make decisions, is accompanied by narrowed spans of control and elongation of hierarchy (Blau, 1968; Meyer, 1968), limiting the number of persons having access to decision makers and reducing aggregate participation. Participation and decentralization are not always in opposition; more often than not they are compatible. However, one cannot determine decentralization by aggregating individual responses concerning participation, since decentralization refers to the position of the person or persons who make decisions

while participation measures the number of people contributing to decisions.

Lincoln and Zeitz introduce further inconsistencies in their interpretation of discrepant empirical results at individual and aggregate levels. The authors find, at the aggregate level, that professionalization decreases administrative intensity, but the opposite relation obtains at the individual level. They speculate that the negative effect of professionalization upon administrative intensity occurs because professionals require less supervision, but that professionalization increases the likelihood of individual persons holding administrative status due to the "expertise or orientation to professional reference groups" (p. 399). In other words, different unmeasured variables—efficiency at the aggregate level but professional socialization at the individual level—are introduced in order to explain discrepant empirical results at the two levels of analysis. This may be informative for purposes of exploration, but it is hardly parsimonious and is inconclusive in the absence of direct measurement of the explanatory variables. Furthermore, it is not always substantively meaningful to analyze simultaneously individual- and organizational-level effects. For example, a relationship between environmental change and administrative intensity is hypothesized at both aggregate and individual levels. At the aggregate level, the conventional argument is made: greater change or uncertainty requires rapid adaptation and hence impedes efficiency under conditions of close supervision. (In fact, the opposite relationship obtained empirically.) At the individual level, Lincoln and Zeitz proffer no explanation for the hypothesized relationship of uncertainty to administration, since the proposition that the environmental instability one perceives determines whether or not he holds an administrative position is illogical. It may be that persons in administrative positions perceive greater uncertainty than others, but the proposition is wholly different from the one tested and cannot be estimated using the recursive model employed by Lincoln and Zeitz.

In sum, Lincoln and Zeitz have misconstrued the meaning of decentralization, have invoked different unmeasured explanatory variables for discrepant empirical results at individual and aggregate levels, and have separated individual from organizational levels of analysis when it was not substantively meaningful to do so.

METHODOLOGICAL ISSUES

Several methodological features of the Lincoln and Zeitz article are of concern. One is their use of Hauser's (1971) covariance model for analyzing contextual effects. The reader's attention is directed toward the top portion of Figure 2 in Lincoln and Zeitz's article. Here, the relationship between aggregate properties ($\bar{X}_j$ and $\bar{Y}_j$) is decomposed into two paths, one of which is constrained to the same regression parameter (b_{yx}) as the corresponding relationship at the individual level, and the other of which is the difference between the aggregate- and individual-level parameters ($b_{\bar{y}\bar{x}} - b_{yx}$). The interpretation of this latter path is at issue here. First, we question whether the path linking the aggregate independent variable and the residual, estimated as the difference between aggregate- and individual-level regression coefficients, is necessarily a causal path. It is not treated as a causal path in Hauser's formulation of the covariance model. The presence of size, which is an exogenous variable, does not warrant treating the ($b_{\bar{y}\bar{x}} - b_{yx}$) path as causal. Second, it is not clear that the residual necessarily reflects a contextual or organizational-level effect. It could be an effect of some unmeasured individual-level variable. Indeed, given the frequency with which the authors invoke unmeasured individual variables in explaining observed effects, they must consider the likelihood that some of the aggregate-level effects they claim are due to misspecification of their model at the individual level.

A further concern is the criterion for statistical significance adopted by Lincoln and Zeitz. There is a 30% likelihood that a coefficient only equal to its standard error is in fact zero. Had the conventional criterion for significance been used, many of the claimed results would have been vitiated. At the organizational level, a negative path from professionalization to administrative intensity and a positive path from administrative intensity to horizontal communication would remain. At the individual level, there would remain only a positive path from administrative intensity to decentralization. If there had been no separation of individual from organizational levels of analysis and the conventional criterion for significance had been used, positive paths from administrative intensity to horizontal communication, communication to decentralization, and professionalization to decentralization would have been reported. This is illustrated in Figure 1. We leave it to the reader to compare this relatively simple model of decentralization in or-

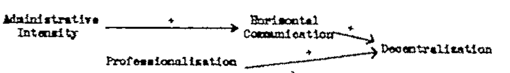


Figure 1: Model of Decentralization in Organizations Following Lincoln and Zeitz

ganizations with the less parsimonious model of Lincoln and Zeitz.

CONCLUSIONS

The Lincoln and Zeitz article is useful in several ways not anticipated by the authors. It demonstrates that some variables describing individuals cannot be meaningfully aggregated to the organizational level and that some organizational variables cannot be meaningfully disaggregated to individuals. It demonstrates that even where equivalences exist between individual- and aggregate-level properties, causal laws need not be isomorphic across levels. Finally, it demonstrates the importance of fully specified theoretical models when applying complex estimation techniques to organizational data.

REFERENCES

Blau, Peter M.
1968 "The hierarchy of authority in organizations." American Journal of Sociology 73:453–67.
Hannan, Michael T., John Freeman, and John W. Meyer
1976 "Specification of models for organizational effectiveness." American Sociological Review 41:136–43.
Hauser, Robert M.
1971 Socioeconomic Background and Educational Performance. Washington, D.C.: Arnold and Caroline Rose Monograph Series of the American Sociological Association.
Meyer, Marshall W.
1968 "Two authority structures of bureaucratic administration." Administrative Science Quarterly 13:211–28.
Williamson, Oliver E.
1975 Markets and Hierarchies. New York: Free Press.

LEVELS OF CAUSATION AND LEVELS OF DATA COLLECTION IN THE STUDY OF ORGANIZATIONAL PROPERTIES*

(REPLY TO MEYER AND STEVENSON)

JAMES R. LINCOLN
Indiana University, Bloomington

GERALD ZEITZ
SUNY, Stony Brook

Sociologists since Durkheim have found it meaningful to distinguish between the global or "structural" properties of social collectivities and those derived from aggregates of individual traits. A key theme of our 1980 ASR paper was that the critical issue regarding collective properties was less the level at which they were first measured (individual versus organizational) than the level of the causal mechanisms generating covariations among them. This issue is at the core of our disagreement with Meyer and Stevenson. Their general stance appears to be that the level of data collection determines the level at which analysis and explanation should proceed. They either fail to understand or deny the validity of our procedures for decomposing covariations among organizational properties into components associated with individual and organizational effects. While we find little substance to their remarks, we appreciate the opportunity to review and expand on certain points in our article.

Meyer and Stevenson's first complaint is that our measure of average membership-share in decisions reflects participation, not decentralization, since the latter concept refers solely to organizational decisions being made at administrative levels lower than the "boss." Yet their claim to insight into the true meaning of decentralization is without foundation either in organizational theory or common sense. An organization which permits a high average level of participation in decisions is decentralized in the fundamental sense of democratization: power is not concentrated in few hands but is widely dispersed. Our article acknowledged that Meyer and Stevenson's notion of decentralization as authority delegation need not vary positively with the level of participation in decision making. But both these (and other) meanings find ample support in the literature, and there is no basis for the claim that one warrants the decentralization label more than another. In attacking our indicator, moreover, Meyer and Stevenson take on an established tradition of empirical organizational research in which this and related aggregate measures of decision-making structure have played prominent roles (see citations in Lincoln and Zeitz, 1980:391).

What is more telling in Meyer and Stevenson's concern with our decentralization measure is their view that it: ". . . describes individual persons as do most of the independent variables in [our] study." They assert that data collected from individuals are useful only for tests of individual-level theories, even, apparently, when the data are gathered with the intent of measuring organizational traits. In this, they oppose the entire thrust of our article. But they have difficulty holding sensibly to this position. They refer to our measure as the number of persons participating in decisions. Whether this is decentralization, our measure of it, or something else, it is a perfectly meaningful organizational property, and it makes no sense at all to hold that it cannot be the object

* Direct all correspondence to James R. Lincoln, Department of Sociology, Indiana University, Bloomington, IN 47401.

of an organizational investigation simply because the data originated with a survey of individual members. Yet we are most disconcerted by their warning that to aggregate data on individuals and perform an organizational analysis runs the risk of introducing "aggregation bias" into model parameter estimates. It is precisely the presence of such "bias" that testifies to the inadequacy of an individual theory to explain variable relationships at the level of the organization. Our paper was explicit in stressing that what is aggregation bias to the analyst aiming to generalize to individuals from observations on groups is to the student of organization solid testimony to the appropriateness of theories positing organizational causal mechanisms.

Further, we did not, in fact, argue for a "negative effect of professionalization upon administrative intensity . . . because professionals require less supervision." This is an unwarranted individual-level interpretation of our hypothesis that: ". . . professionalism and instability are both forces rendering close supervision and high administrative overhead ill-suited to effective performance" (Lincoln and Zeitz, 1980:398). Also, their assertion that our hypotheses depend on unmeasured variables is true enough but trivial, as any model assumes unmeasured intervening processes and is no less "parsimonious or conclusive" because of it. Moreover, they again miss the point: the causal processes relating professionalism and administration are presumed to be different at organizational and individual levels. That is precisely why it is of interest to separate the two empirically.

Meyer and Stevenson next suggest that it is not meaningful to separate individual and organizational effects via our methodology when the specification of cause and effect reverses over levels of aggregation. They state that we ". . . proffer no explanation for the hypothesized relationship of uncertainty to administration, since the proposition that the environmental instability one perceives determines whether or not he holds an administrative position is illogical." That proposition is not logically flawed, but we did, in fact, advance the same prediction that Meyer and Stevenson adopt as their own: that at the individual level perceptions of uncertainty follow from occupancy of administrative position rather than the other way around. While we reject their conclusion that a decomposition analysis of individual and structural effects is thereby ruled out, we did acknowledge the modeling complications that this circumstance creates. The statistical issue is that in partitioning the total aggregate effect ($b_{\overline{Y}\overline{X}}$) into its within ($b_{yx}$) and adjusted between ($b_{\overline{Y}\overline{X}} - b_{yx}$) components, b_{yx} (and therefore $b_{\overline{Y}\overline{X}}$ as well) is a Biased estimate if the true causal parameter in the individual model is b_{xy}.

This is an interesting problem, and we would certainly invite other views on it. Our published position was that "Since our theoretical model is specified at the organization level, the causal ordering [of uncertainty antecedent to administrative intensity] we first stipulated seems the appropriate one to retain" (Lincoln and Zeitz, 1980:399). In hypothesizing a relationship between organizational properties, in other words, one has a theory that defines the level at which the effect occurs. In our case, that theory placed the locus of causation at the organization level. Our hypothesis proved right about the level of the effect but wrong about its sign, which was positive instead of negative. Virtually no part of the significant aggregate effect could be described by the within-group slope. Had an individual-level process proved dominant, we might have reasoned as follows. There is evidence in the data that the aggregate relation is due to an individual process. This is incompatible with the theory relating environmental change to bureaucratic forms. The most plausible individual-level alternative theory specifies the opposite causation, from administrative position to perceived environmental change. We should therefore respecify the model in accordance with this theory and reestimate it. In other words, one's organizational theory should fix the level of the process believed to generate a relationship between organizational properties. Evidence for the causal priority of a different level demands rejection of the initial theory and revision of its causal assumptions.

Meyer and Stevenson are further troubled by what they see to be our causal interpretation of the dependence of the adjusted aggregate dependent variable $[R=\overline{Y}_j-b_{yx}(\overline{X}_j-\overline{X})]$ on the aggregate independent variable, $\overline{X}_j$. This coefficient, $b_{R\overline{X}}$, is equal to the difference in the between and within-group regression slopes, $b_{\overline{Y}\overline{X}}-b_{yx}$. They fail to explain why it should not be so interpreted except to note that Hauser (1971:17) specifies no corresponding causal path in his diagram of the ANCOVA model. It is true that Hauser did not elaborate the ANCOVA algebra as far as we. Except for our specification of a global variable, S (which, by the way, we never implied had any bearing on the interpretation of $b_{R\overline{X}}$), our model is identical to that discussed by Duncan et al. (1972) and Alwin (1976). We primarily credited Hauser because his was the first published attempt to model ANCOVA in this way, and his discussion remains one of the most detailed.

Meyer and Stevenson's cryptic claim that $b_{\overline{Y}\overline{X}}-b_{yx}$ has no causal significance but merely reflects unmeasured individual-level variables

also echoes Hauser's (1970, 1971) critique of contextual analysis. Hauser's important discussion of this topic has too often been taken as a blanket denunciation of multilevel analysis, much as Robinson's (1950) paper raised the spectre of ecological fallacy which still haunts naive researchers putting aggregate data to quite legitimate use. His critique, however, was subtle and largely specific to a then-fashionable approach to a narrow set of research problems. Hauser did indeed argue that apparent group effects in models of individual-level outcomes may be due to poor specification of individual-level causes. But read in context this lends no authority to Meyer and Stevenson's unsubstantiated charge that the structural effect parameter in our aggregate model is intrinsically misspecified because of omitted individual variables. Hauser's concerns were chiefly with those "school effects" studies which sought to detect relationships between individual student achievement and school averages on a variety of student background characteristics. When such associations proved statistically significant in a body of data, investigators were quick to attach post hoc social structural interpretations. No attempt was made to measure such group properties directly, and the aggregate indicators were typically far removed from the properties they were said to represent. However, Hauser (1970:661) anticipated that forms of contextual analysis might legitimately be applied to problems in the interpretation of collective properties, for while he stated that ". . . it is not necessary to control for individual attributes when you are really interested in the covariation of structural or aggregate variables," he observed that "Such controls may be desirable when the object is to find out how structural or aggregate variation occurs, but this is not ordinarily the case in contextual analysis."

That indeed has not been the case in most contextual analyses, for, as our article noted, such studies are typically based on individuals as units of analysis. However, "to find out how . . . aggregate variation occurs" aptly describes the goals of our work. We discussed a research tradition in which investigators have measured and analyzed properties of organizations with data aggregated from individual members' reports. However dubious their validity in occasional studies, such measures were designed to address organizational concepts. We drew from received organizational theories in proposing a model which defined a set of causal dependencies on those organizational properties. That model was no doubt misspecified in certain respects; social science models often are. But Meyer and Stevenson's comment is irrelevant, for it applies to our analysis a criticism made of work which has very little in common with our own. In estimating each total aggregate regression slope, $b_{\bar{Y}\bar{X}}$, we assume that $\bar{Y}_j$ is causally dependent upon $\bar{X}_j$. Once that assumption is made, an inquiry into whether and how the relation combines processes at organizational and individual levels demands that attention be directed to the within (b_{yx}) and adjusted between ($b_{\bar{Y}\bar{X}}-b_{\bar{y}x}$) group slopes. If the total aggregate effect is described by the within-group slope, an individual causal mechanism is shown to be operative. The degree to which that parameter fails to describe this covariation is, reasonably enough, indicated by the difference between the total aggregate slope and the average within-group slope. If this difference is large and significant, then the process generating the relation is presumed embedded at a supraindividual level. The point is that, having previously made the specification that $\bar{X}_j$ causes $\bar{Y}_j$, $b_{\bar{Y}\bar{X}}-b_{\bar{y}x}$ speaks directly to the substantive question of whether an organizational mechanism is present, and the issue of whether it is biased is moot.

Meyer and Stevenson's final assertion that .05 is the one true level of statistical significance for rejecting null hypotheses carries roughly the same authority as their previous claim on the true meaning of decentralization. Our analysis was based on an N of 20 organizations. In a sample of this size, $\alpha = .05$ is an extremely conservative test of the significance of a sample estimate. There is wide agreement that higher probabilities of making a Type 1 error are in order when samples are very small (e.g., Heise, 1969:61). In fact, apart from the force of convention, a t-ratio equal to or greater than 1.0 is less arbitrary than 2.0, since, as we (Lincoln and Zeitz, 1980:403) pointed out, it is the cut-off point for determining whether a regressor adds to prediction more than is expected by chance.

CONCLUSIONS

Despite the preoccupation of writers like Meyer and Stevenson with whether one measures collective properties by aggregating survey responses or by some means which generates a single organizational score, the issue has been overplayed. A measured covariation between global indicators of organizational attributes may easily reflect an individual mechanism, just as aggregate associations often stem from "structural effects." The fact that a measure can or cannot be disaggregated has nothing to do with the level of the process driving it. Yet valid aggregate measures of collective properties have an analytic advan-

tage over global measures because they do allow for some limited inquiry into the nature of that process. Moreover, aggregate indicators are probably more reliable than global indicators, since the process of pooling fallible individual scores serves to cancel out random measurement errors. Even such "objective" organizational properties as size are therefore best measured by aggregating the reports of different informants (Seidler, 1974). As we (Lincoln and Zeitz, 1980:393) discussed, to control for individual effects in this circumstance is to control for sources of measurement bias. Clearly, however, these issues in macrosocial measurement and analysis deserve closer and greater attention by social scientists than they have received in the past.

REFERENCES

Alwin, Duane F.
1976 "Assessing school effects: some identities." Sociology of Education 49:249–303.

Duncan, Otis Dudley, David L. Featherman, and Beverly Duncan
1972 Socioeconomic Background and Achievement. New York: Seminar Press.

Hauser, Robert M.
1970 "Context and consex: a cautionary tale." American Journal of Sociology 75:645–65.
1971 Socioeconomic Background and Educational Performance. Washington, D. C.: Rose Monograph Series.

Heise, David R.
1969 "Problems in path analysis and causal inference." Pp. 38–74 in E. F. Borgatta (ed.), Sociological Methodology, 1969. San Francisco: Jossey-Bass.

Lincoln, James R. and Gerald Zeitz
1980 "Organizational properties from aggregate data: separating individual and structural effects." American Sociological Review 45:391–408.

Robinson, W. S.
1950 "Ecological correlations and the behavior of individuals." American Sociological Review 15:351–7.

Seidler, John
1974 "On using informants." American Sociological Review 39:816–31.

American Sociological Review

Volume 46 — December 1981 — Number 6

SCOPE AND MISSION OF AMERICAN SOCIOLOGICAL REVIEW: The *American Sociological Review* publishes work of interest to the discipline in general, new theoretical developments, results of research that advance our understanding of the most fundamental social processes, and important methodological innovations. Like other publications of the Association, emphasis is given to exceptional quality. Unlike the more specialized journals of the Association, the *American Sociological Review's* primary focus is on whatever has the most general bearing on the knowledge of society.

Concerning manuscripts, address: Editor, *American Sociological Review*, Department of Sociology, University of Illinois, Urbana, Illinois 61801. (217) 333-6515.

The *American Sociological Review* (ISSN 0003-1224) is published at 49 Sheridan Avenue Albany, New York, bimonthly in February, April, June, August, October, and December.

Annual membership dues of the American Sociological Association, $20–$120 ($10 of this amount may cover a subscription to the *American Sociological Review*); Associate, $25; International Associate, $25; Student Associate, $12.

Application for membership and payment of dues should be made to the Executive Office. Subscription rate for members, $10; nonmembers, $18; institutions and libraries, $36; nonmember students, $12. Single issue, $4.

New subscriptions and renewals will be entered on a calendar year basis only.

Change of address: Six weeks advance notice to the Executive Office and old address as well as new are necessary for change of subscriber's address.

Concerning advertising, changes of address, and subscriptions, address: Executive Office, American Sociological Association, 1722 N Street, N.W., Washington, D.C. 20036.

Claims for undelivered copies must be made within the month following the regular month of publication. The publishers will supply the missing copies when losses have been sustained in transit and when the reserve stock will permit.

Second class postage paid at Washington, D.C. and additional mailing offices.

(Articles in the REVIEW are indexed in the Social Sciences Index, Psychological Abstracts, Sociological Abstracts, Ayer's Guide, University Microfilms, Abstracts for Social Workers, International Political Science Abstracts, and United States Political Science Documents.)

AMERICAN SOCIOLOGICAL REVIEW

Official Journal of the American Sociological Association

DECEMBER, 1981 VOLUME 46, NUMBER 6

(Continued on next page)

OTHER

NEW EDITOR

Effective with the February 1982 issue, Sheldon Stryker will become the editor of the American Sociological Review. Beginning July 1, 1981, send all manuscripts to

Sheldon Stryker, Editor
American Sociological Review
Institute of Social Research
1022 East Third Street
Bloomington, IN 47405

CONTRIBUTORS TO THIS ISSUE

■ ANNE B. SHLAY (Keeping Up the Neighborhood) is Assistant Professor, Department of Consumer Economics and Housing, Cornell University. She is studying the dynamics of zoning, focusing on the relationship between neighborhood socioeconomic composition and changes in the regulation of land use. She also studies the impact of zoning on changing distributions of property and land values.

■ PETER H. ROSSI is Professor of Sociology and Director of the Social and Demographic Research Institute, University of Massachusetts at Amherst. He is studying methodological advances in factorial surveys, evaluation and applied social research, weapons and violent crime, and estimation of the damages from natural hazards to households and individuals. He is coauthor, with Richard A. Berk and Kenneth J. Lenihan, of *Money, Work and Crime* (Academic Press, forthcoming), with Howard Freeman, of *Evaluation: A Systematic Approach* (Sage, forthcoming), and with J. D. Wright and S. Anderson, of *Handbook of Survey Research* (Academic Press, forthcoming).

■ KAZIMIERZ M. SLOMCZYNSKI (Stratification, Work, and Values) is from the University of Warsaw, Poland, and is Visiting Scientist, National Institute of Mental Health. He is involved in a cross-national analysis of social stratification and value systems and in a dynamic analysis of occupational structure with an emphasis on job careers and educational attainment.

■ JOANNE MILLER is Research Sociologist, Laboratory of Socio-Environmental Studies, National Institute of Mental Health. Her current research includes a comparative analysis of intellectual functioning in Poland and the United States (with Slomczynski and Kohn) and an assessment of receptivity to change in Polish society (with Slomczynski). She is also studying the measurement and meaning of supervision in men's and women's jobs and preparing a review of recent literature on sex roles for the Annual Review of Sociology (with Howard Garrison).

■ MELVIN L. KOHN is Chief, Laboratory of Socio-Environmental Studies, National Institute of Mental Health. He is presently completing a longitudinal inquiry into the interrelationship of social stratification, job conditions, and personality. With Karen A. Miller and Carmi Schooler, he is analyzing the effects of educational experience on personality development.

■ ALDON MORRIS (Black Southern Sit-In Movement) is Assistant Professor, Department of Sociology and the Center for Research on Social Organization, University of Michigan. His research interests are Political Sociology, complex organizations, and social movements. He is studying the social structure of the American Civil Rights movement. He is author of *Origins of Southern Civil Rights Movement* (Free Press, forthcoming), an analysis of the organizational foundation of that movement.

■ JOHN W. GARTRELL (Inequality within Rural Communities of India) is Associate Professor, Department of Sociology, University of Alberta at Edmonton, has a continuing interest in problems of economic growth and inequality as well as the effects of social structure on innovation in agrarian communities in India. He has also recently completed interview surveys investigating the social impact of rapid change in several communities in the oil sands region of northern Alberta.

■ GARY KLECK (Racial Discrimination in Criminal Sentencing) is Assistant Professor, School of Criminology, Florida State University. He is studying gun control, the subculture of violence, and the use of self-reporting methods in determining the class distribution of crime.

■ RICHARD A. BERK (A Test of the Stability of Punishment Hypothesis) is Professor, Social Process Research Institute and the Department of Sociology, University of California at Santa Barbara. He is involved in an extension of the research reported in this issue and especially formulating dynamic models of "punishment" to account for historical patterns in California. He is starting a three-year research project on wife battery and completing a study of citizen evaluations of the performance of various aspects of the criminal justice system. With Peter H. Rossi and Kenneth J. Lenihan, he is coauthor of *Money, Work and Crime* (Academic, forthcoming), and with Thomas Cooley and Charles J. LaCivita, coauthor of *Water Shortages* (Abt Books, forthcoming).

■ DAVID RAUMA is a Ph.D. candidate, Department of Sociology, University of California, Santa Barbara and is studying the impact of transfer payments on the post-prison experiences of released felons and the causes and consequences of family violence. He is also involved in a quantitative analysis of historical patterns in the use of punishment in California.

■ SHELDON L. MESSINGER is Professor, School of Law, University of California at Berkeley. With Andrew von Hirsch and Richard Sparks, he is studying strategies of determinate sentencing and, with Richard Berk, long-term trends in imprisonment. He is coeditor, with Egon Bittner, of *Criminology Review Yearbook, Vol. II* (Sage, 1980).

■ THOMAS F. COOLEY is Professor, Department of Economics, University of California at Santa Barbara. He is studying identification theory in econometrics. He is coauthor of *Water Shortages* (Abt Books, forthcoming) with Richard A. Berk and Charles J. LaCivita.

■ DARRELL MONTERO (The Japanese Americans) is Associate Professor and Director, Urban Ethnic Research Program, Arizona State University. He has completed a national study of Vietnamese refugees and their socioeconomic adaptation to the United States. His current research focuses upon ethnic communities, including Japanese, Vietnamese and Cuban.

■ ROBERTO M. FERNANDEZ is Ph.D. candidate, Department of Sociology, University of Chicago. His dissertation topic is "Job Contacts and Labor Market Structure." With François Nielsen, he studies bilingualism and school achievement.

■ JANE C. KULIK is a graduate student in Sociology, University of Chicago and of Abt Associates. She is evaluating social program implementation and program effectiveness, specifically recent welfare reform demonstrations.

■ BETH MINTZ (Interlocking Directorates and Interest Group Formation) is Assistant Professor, Department of Sociology, University of Vermont. She is studying interlocking directorates and class cohesion. She is coauthor, with Michael Schwartz, of *The Structure of Power in the American Corporate System* (forthcoming).

■ MICHAEL SCHWARTZ is Associate Professor of Sociology, State University of New York at Stony Brook. His recent work has been on intercorporate power and coordination, and includes two forthcoming books, *The Structure of Power in the American Corporate System* (with Beth Mintz) and *The Business as a Ruling Class* (with 12 coauthors).

■ LYNNE G. ZUCKER (Taxonomies of Institutional Structure) is Assistant Professor, Department of Sociology, University of California at Los Angeles. She is investigating stability and change in institutional structures. She is author of *Organizations As Institutions: Restructuring the Social Order* (forthcoming).

■ CAROLYN ROSENSTEIN is a doctoral student, Department of Sociology, University of California at Los Angeles. She plans to do her dissertation on the continuity and/or discontinuity of work of immigrants.

■ ROBERT NASH PARKER (Structural Constraints and Individual Career Earnings Patterns) is Assistant Professor, Department of Sociology, University of Akron. His current research interests include a continued investigation of types of homicide, including the impact of levels of aggregation, measurement error, and ratio variables on the results of such studies. He is also working on the application of various statistical techniques from psychology, sociology, and econometrics to life history data.

■ MICHAEL E. SOBEL (Diagonal Mobility Models) is Research Sociologist, Office of Research and Evaluation, U.S. Bureau of Labor Statistics. His current projects include a cross-national comparative analysis of occupational mobility, and an analysis of occupational mobility in 19th century America. He is author of *Lifestyle and Social Structure* (Academic Press, forthcoming).

■ G. EDWARD STEPHAN (The Historical Distribution of County Seats) is Professor, Department of Sociology, Western Washington University. He is studying time-minimization theory development and the consequences of deviation from time-minimization laws.

■ DOUGLAS R. McMULLIN is a graduate student, Department of Sociology, Western Washington University. His M.A. thesis is about nations which deviate from the size-density law.

■ MICHAEL L. RADELET (Racial Characteristics and the Death Penalty) is Assistant Professor, Department of Sociology, University of Florida. He is studying racial correlates of disparities between official police reports and court records of Florida homicides and racial correlates of decisions made by the Florida Supreme Court. He is also working on the utilization of health services for problems related to psychosocial distress.

KEEPING UP THE NEIGHBORHOOD: ESTIMATING NET EFFECTS OF ZONING*

ANNE B. SHLAY
Cornell University

PETER H. ROSSI
University of Massachusetts

A political dimension is added to a framework of urban ecology by examining the impact of zoning on housing and population growth within the Chicago metropolitan area. Looking at the social and demographic changes within a sample of 395 central city and suburban tracts, the analyses model the effects of zoning on growth from 1960 to 1970. Findings show that by regulating the types of housing that may be included within local areas, zoning affects local housing availability and hence the socioeconomic composition of sub-areas in the Chicago SMSA. Such findings suggest that the urban ecology framework requires modification to account for the influence of political processes.

INTRODUCTION

Urban ecology seeks to portray a dynamic picture of urban social change. Asking how people organize and adapt themselves to the physical contraints within an urban setting, the ecologist examines the mechanisms through which spatial organization is created and maintained within the metropolis (Hawley, 1950). Specifically, an ecological focus looks at the relationships among positions in the urban space (locations), positions in the urban social structure (class, race, occupation, education, etc.) and the distribution of land use (Park, Burgess, and McKenzie, 1967; Schwirian, 1974a; Berry and Kasarda, 1977). Hence, the mechanisms which sort people into so-called "natural areas" explain the way in which a social structure is reproduced within urban society.

The social structure to which urban ecology traditionally has been attentive is apolitical. In the 1920s, when ecological frameworks originated, local government made few, if any, direct attempts to regulate the uses of land. Some political actions taken in that period affected land use and growth patterns—transportation franchises were given to entrepreneurs and street patterns were laid out—but today's complex of land use regulations and building codes were in embryonic stages. Unregulated land use characterized the Chicago upon which Park and Burgess superimposed concentric zones centered around the Loop, and at least some of today's ecologists persist in slighting the importance of governmental regulations in urban development and the maintenance of land use patterns.

For example, Schwirian's (1974b) collection of writings on urban ecology provides only scanty discussions of the politics surrounding the urban planning process. Berry and Kasarda (1977) appear to dismiss planning as politics within the United States, arguing that in contrast to nation- or state-wide planning, localized planning decisions reflect competitive market processes and hence are outcomes rather than inputs into models of urban growth and development.

To be sure, we are not the first to note

* Direct all correspondence to Anne B. Shlay, Department of Consumer Economics and Housing, Cornell University, Martha Van Rensselaer Hall, Ithaca, NY 14853.

Prepared under Grant Number HUD 2367 from the Department of Housing and Urban Development. We are indebted to Jeffrey K. Liker, Lisa Sushelsky, Eleanor Weber-Burdin, and James D. Wright for their devoted help at crucial points, to Bill Form for wise and incisive editorial guidance, to Sabella J. Bradley for craft-like work with the typewriter, and to two anonymous reviewers.

the apolitical character of the urban ecological framework. We follow Form (1954), Molotch (1976), and Logan (1976, 1978, 1980) in arguing that communities self-consciously regulate local land use and therefore may not simply respond to the microeconomic competitive determinants posed by traditional ecology. With zoning, as Form argues, a political dimension is inserted into those mechanisms influencing ecological maintenance and change (Form, 1954). Metropolitan social and spatial organization may in part result from competition between local communities for what they regard as more advantageous land use (Molotch, 1976; Logan, 1976). If, as Logan and Molotch argue, political institutions are central to metropolitan social-spatial structure, then outcomes of local political activity—in particular, zoning—should directly influence changes in metropolitan ecological patterns. Rather than look at the process of political negotiations over land-use decisions (Willhelm, 1962), the motivations for land-use regulations (Molotch, 1976), the social organization of land-use actors (Form, 1954), or the potential for political activity based on community social structure (Logan, 1976, 1978, 1980), we explicitly look at the ways in which zoning affects metropolis-wide neighborhood change. This paper provides some empirical evidence on how the political decisions involved in zoning affect urban growth patterns.

Early ecologists described "natural" distributions of populations generated by competition, invasion, succession, etc., because these were the factors that clearly operated to determine the spatial arrangements during a period of unregulated industrial and urban development. In more contemporary times the politics which regulate land use may modify or even take precedence over "natural" unfettered competition and succesion.

The relative absence of politically distributed, area-based goods and services during ecology's formative stages also influences other aspects of ecological theory. For the traditional ecologist, the value of a particular location was based on its accessibility vis á vis the central business district (Hawley, 1950), not on characteristics associated with the neighborhood itself. Hence, ecological models did not address the full consequences of location. Without denying the importance of accessibility to the central business district, we follow others in suggesting that urban locations provide differing access to public and private goods and services (Harvey, 1973; Molotch, 1976; Logan, 1978). According to Harvey (1973), "Inherent in social processes lies the question of spatial organization. Externality effects are localized, so are job and housing opportunities, resource benefits, communication links, etc. Political power is partly arealy based. *Many of the hidden mechanisms for redistributing income come to fruition in the act of location*." (Emphasis added.)

This study takes one major part of the local political process—zoning—and looks at its impact on metropolitan development by viewing zoning as an important determinant of land use. We do not suggest that zoning regulations have replaced market mechanisms, but only that such processes are modulated by zoning in substantively significant ways.

ZONING AND EXCLUSION

Zoning as a means to insure "a place for everything and everything in its place" is an accepted practice (Perin, 1977). A legal process which reserves land parcels for specifically designated uses, zoning can segregate single family homes from apartments, restrict industry to particular locations, and create special areas for shopping or commercial purposes. Zoning also often regulates the size of structures and building lots, through "bulk and area" requirements that typically stipulate maximum building height and minimum dimensions for the building-lot area, lot frontage (width), floor area, and front, side, and rear yards. Such requirements are often specified for each type of land use, to the point of regulating the minimum amount of land required per bedroom.

Over half a century ago, zoning first was proposed as a method to rid New York City's Fifth Avenue luxury retail establishments of adjacent garment fac-

tory lofts. At the same time, housing reformers saw zoning as a means to improve housing for the poor (Bassett, 1936; Friedman, 1968; Toll, 1969; Fitch, 1977). Liberal reformers and upper-class conservatives found common cause in zoning, since the aims of both were easily subsumed under the broad issue of "land-use incompatability." Convinced that the mixing of factories with residences was detrimental to families' "health, safety, and welfare," in 1926 the Supreme Court ruled that zoning was constitutional under the broad mandate of police power (Euclid v. Ambler Realty Co. 272, U.S. 365). In essence, local power to control land use for the general good of family life was declared to be more important than the specific interests of individual property owners.

Because defining zoning as a broad welfare measure without planning direction would lead to arbitrary and eclectic local definitions of land-use incompatability, the court defined acceptable zoning as following a specific scheme of land-use priorities, the "Euclidean Pyramid" (named after the case in question).

Euclidean zoning designated a rank ordering of land-use desirability to guide local governments in their decisions. Zoning's primary purpose was specified as protecting neighborhoods of single-family dwellings, placed at the top of the priority pyramid. Single-family land use is followed by more "permissive" residential land uses, e.g., duplexes, apartment houses, etc. The increasing width of the pyramid indicates a successive accumulation of land uses with greater permissiveness. Thus, at the base of the pyramid is industrial zoning, which permits all types of land use.

Euclidean zoning's definition of "good neighborhoods" (single-family housing) versus "bad neighborhoods" (mixed land use) is strengthened by a number of additional rationales. First, as a "protectionist device," zoning is justified as a public welfare measure. Zoning's benign intent is to protect residential neighborhoods from the congestion, noise, traffic, pollution, and general ugliness associated with commerce and industry. Zoning seeks to protect the rights of all individuals by segregating land use in a manner that maximizes the well-being of the entire community (Friedman, 1968; Williams, 1975).

Secondly, zoning is designed to "protect" property values on behalf of home owners. By segregating perceived deleterious land use, zoning acts as a brake upon market forces. As an economic policy, zoning distributes land use so that property will have maximum value on the market (Babcock, 1969).

Thirdly, zoning works to "internalize externalities," that is, to segregate "lower" types of land use (industry, commerce, and apartments) and their accompanying nuisances from the "highest" type of land use (single family homes). In this way, the noxious side effects of commercial and industrial uses are negated because they are kept within specifically designated areas where those land users who generate such externalities are constrained to bear them.

Finally, as a method to protect the public's welfare and, at the same time, to protect property values, zoning becomes a local planning policy. As a "tool of planning" (Haar, 1959; Williams, 1966; Delafons, 1969; Babcock, 1969), zoning is juridically and practically justified as a method to be used "in accordance with a comprehensive plan" (Haar, 1959). By combining land-use regulations with comprehensive planning, zoning is intended to guide land use in accordance with perceived, long-range local needs.

With experience, these principles have also become tied to the less abstract goal of managing local tax revenues (Williams and Norman, 1971). Through such "fiscal zoning," local government can limit per capita expenditures and also can expand local tax bases. To preserve the taxable value of land while limiting the tax rate, a community may exclude potential negative externalities and simultaneously fence out land users who might not "pay their way" (Babcock, 1969; Williams and Norman, 1971; Danielson, 1976; Moskowitz, 1977).

Of course zoning affects more than the mere flow of money in and out of city hall. In striving to achieve fiscal goals, a community necessarily regulates its socioeco-

nomic composition. In this way, fiscal zoning is also a social planning policy that may be directed at or result in socioeconomic exclusion. Termed "exclusionary zoning," such practices indirectly operate as methods for legal socioeconomic and racial discrimination. Therefore, zoning to maximize a community's public welfare simultaneously works to maintain a community's residential status quo (Downs, 1970, 1973; Davidoff and Davidoff, 1971; Danielson, 1976; Bergman, 1974). In other words, zoning is a method for "keeping up the neighborhood."

Zoning is purported to achieve desired distributions of land use by restricting the free competition among potential bidders for land parcels. If zoning is enforced, vacant land specified for single family homes will not be bought or developed by persons or corporations who intend to develop high-density land use or nonresidential establishments. In addition, already developed parcels under existing zoning regulations may not be altered without special political exceptions. Zoning, however, is not fixed for eternity. It can be nibbled away by the granting of variances until exceptions destroy the rule. Zones may be changed altogether, as when land zoned for agricultural use may be reclassified to some restrictive residential classification. Indeed, it is the processes of zoning changes and, in particular, variances that have led some observers to conclude that market forces cannot be resisted by the politics of zoning, at least in the long run (Neutze, 1968; Danielson, 1976; Nelson, 1977).

There are many issues in the understanding of the impacts of zoning on land use that cry out for empirical verification. We cannot address all such issues in the restricted compass of one article. Rather we will concentrate on the issue of how zoning affects the patterns of growth (or decline) within local neighborhoods. If zoning has effects, then we should expect that the zoning in place in one point in time should affect what happens in that locality over some ensuing period. Thus land parcels zoned for single-family homes in 1960 should evidence different patterns of growth over a decade than sites allowing more varied land uses. This line of inquiry does not address itself to the issue of what forces produce differentiated zoning; rather, it starts with a given pattern of zoned land use and asks about the consequences of such patterns.

The political processes that created zoning ordinances have been discussed by Willhelm (1962) and Makielski (1966). Our concern is with the consequences of zoning once zoning ordinances are in place.

DESIGN AND DATA BASE

Our design links zoning classifications in force at the beginning of the study period (1960) with changes in housing and population stocks over the ensuing decade for a sample of census tracts drawn from the Chicago SMSA. Census tracts are selected as units of analysis both because they can serve as proxies for neighborhoods and because they are areal units sufficiently small to be relatively homogeneous in zoning classifications applied.[1]

The general model employed in the analysis is as follows:

$$T70_{ij} = f(T60_{ij}, \Sigma Z60_{ik}, \Sigma C60_{im})$$

where

$T70_{ij}$ = the 1970 Census value of some housing or population variable j for some tract i

$T60_{ij}$ = the 1960 Census value of some housing or population variable j for tract i

$\Sigma Z60_{ik}$ = vector of k 1960 zoning classifications for tract i

$\Sigma C60_{im}$ = vector of 1960 m characteristics of tract i

---

[1] Because tracts may change in population size, the Census Bureau redefines some tract boundaries at each census taking. These tracts are changed in two ways. First, a 1960 tract can be split into two, three, or more tracts in 1970. These split tracts could be added together in order to form a comparable area with their respective 1960 counterparts. Second, tract boundaries can be shifted, resulting in areas that are not comparable over time. While such boundary changes are usually quite minor, some are quite substantial. Therefore, we chose to include only perfect one-to-one tracts and perfect combinations of tracts and to drop tracts which did not form perfectly comparable areas.

The model assumes that the best predictor of the status of a tract's value in some respect (e.g., the number of housing units in that tract) is the number of such units in 1960 modified in some way by the zoning restrictions in place for the tract in 1960 and certain other characteristics of the tract in 1960 (e.g., the amount of land area, tract location in the SMSA, etc.).

It should be noted that the effects of zoning are incorporated into the model in two ways: directly in the form of zoning classification variables reflecting the regulations in force in 1960, and indirectly through the historical effects of zoning that are embodied in the status of the tract in 1960.

The function form given to the model in this paper is OLS regression. Since we have no theoretical reason to prefer one form over any other, we have chosen to use the most convenient, OLS.

This study relies on two data sources: first, the zoning regulations of the city of Chicago and of selected suburbs, cities, and counties within the Chicago SMSA that were in force in 1960; and second, census tract data from the 1960 and 1970 Censuses.

The two-stage sample of census tracts was drawn as follows: first, the Chicago SMSA was divided into two parts, the city of Chicago and the rest of the SMSA. At the second stage, central city tracts were stratified by two characteristics: housing density and mean household-income levels. Tracts were sampled within strata in order to maximize variance in zoning.[2] In selecting suburban tracts, we sampled political subdivisions in order to maximize the zoning variation among political units. Suburban municipalities were stratified according to the total number of tracts within each political unit: municipalities containing one tract, two to four tracts, five to nine tracts, and ten or more tracts. Municipalities were randomly selected from each strata so that the suburban sample consisted of the distribution of tracts contained within these political units.

The zoning regulations for sampled tracts were obtained by site visits to the zoning agencies for the political subdivisions involved. Eliminating tracts for which no census data and/or zoning data could be obtained left a final sample of 395 census tracts: 249 within the central city and 146 tracts shared among 39 local political subdivisions and 2 counties (See Appendix).[3]

Measuring Zoning

Land-use regulation in Illinois is left to local communities. As a consequence, local zoning regulations vary from place to place, although, fortunately, overall Euclidean pyramid principles prevail.

To make local residential zoning districts comparable across communities, land-use categories were grouped by the two most common rules used: permitted land uses within each district and minimum lot areas per dwelling unit. The resulting eleven standardized residential zoning categories expectedly are ordered according to the Euclidean hierarchy in which single-family dwellings are the "highest and best use" and low densities are "higher" than high density regulations.

The use and area requirements defining each standard zoning level are shown in Table 1. All single-family dwelling categories permit only single family dwellings, all two-family dwelling categories permit both duplexes and single-family dwellings, and all multiple-family dwelling categories permit all types of residential land use. Hence, these levels range from the most restrictive zoning category, which permits only single-family dwellings on large lots, to the most permissive zoning category,

[2] Since Chicago tracts tended to be concentrated in more permissive zoning classifications, a straight random sample of tracts would have produced an uneven distribution of zoning classifications within the city and hence restrict the variance in zoning among tracts. The sampling strategy employed tended to achieve a more even distribution of zoning and hence a larger variance in zoning practices than is representative of the city as a whole. In terms of the regression analyses used, this sampling strategy produces unbiased regression coefficients, but lowers sampling errors and hence enhances the capabilities of rejecting Type II errors.

[3] Unincorporated areas were governed by county zoning agencies.

Table 1. Standard Residential Zoning Levels Code

Symbol	Minimum Lot Area[a]	Weight Used
Single-Family-Dwelling Zoning Categories		
S-1	43,561 to 217,800 sq. ft.	11
S-2	21,800 to 43,560 sq. ft.	10
S-3	10,001 to 21,799 sq. ft.	9
S-4	7,201 to 10,000 sq. ft.	8
S-5	0 to 7,200 sq. ft.[b]	7
Two-Family-Dwelling Zoning Categories		
T-1	3,000 to 6,500 sq. ft.	6
T-2	0 to 2,999 sq. ft.[b]	5
Multiple-Family-Dwelling Zoning Categories		
M-1	2,001 to 4,000 sq. ft.	4
M-2	1,001 to 2,000 sq. ft.	3
M-3	501 to 1,000 sq. ft.	2
M-4	0 to 500 sq. ft.[b]	1

[a] Minimum lot area per dwelling unit for the primary permitted land use within each category (e.g., minimum lot area per dwelling unit within a multifamily structure).

[b] "0" does not mean that no lot is required, but rather that no minimum lot area is specified.

which permits any type of residential land use on very small lots.[4]

To use these residential zoning levels, the zoning classifications used in each local community were examined in detail and each residential category was placed in the "Standard" coding category closest to the local definition. No difficulty was found in coding residential zoning: of the 42 communities studied, all of the residential zoning categories could be fitted into the standard scheme shown in Table 1.

The Euclidean scheme also made it possible to make standard nonresidential zoning districts that reflect degrees of undesirable land use. The seven standard nonresidential zoning districts are ranked according to potential negative externalities: Undeveloped Land, Office-Research, Limited Retail Trade, General Retail Trade, General Commercial, Light Manufacturing, and General Manufacturing.

The overall residential zoning status of each tract was measured by mapping tract boundaries on zoning maps. The percentage of tract area classified into each of the standard zoning levels shown in Table 1 in each tract was calculated.[5] Using the weights shown in Table 1, an "average" residential zoning level was computed for each tract and defined as the Exclusionary Zoning Index. For each tract, the percentage zoned for each level of residential zoning restrictiveness (i.e., S-1, S-2, etc.) was multiplied by its "weight" (S-1 = 11, S-2 = 10, . . ., M-4 = 1) relative to other categories in the hierarchy. These weighted percentages were combined into a composite score and divided by the total percentage of each tract zoned for residential land use:

$$\text{Exclusionary Zoning Index} = \Sigma_i W_i P_i / \Sigma_i P_i$$

where W is the weight assigned to land-use category i, and P_i is the percentage of the tract zoned for land-use category i. The Exclusionary Zoning Index ranges from 0 to 11, a low score representing tracts containing largely permissively zoned residential land and a high score representing tracts containing quite restrictively zoned residential land. In addition, for each tract, we calculated the percentages of area zoned for business and manufacturing use.

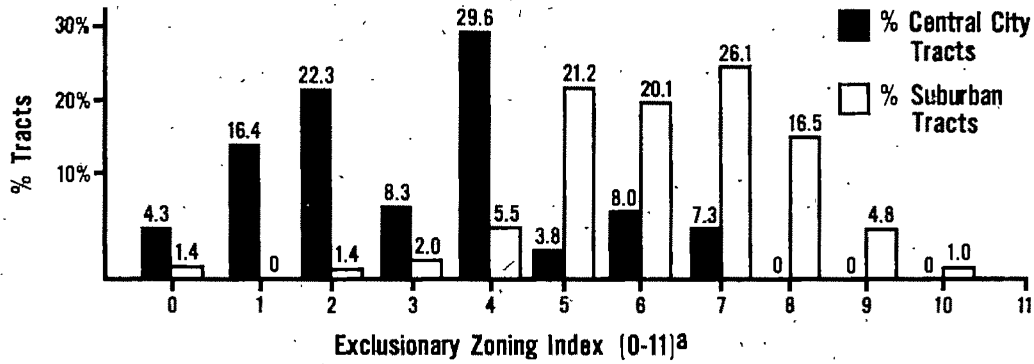
The distribution of zoning among central city and suburban tracts in 1960 is shown in Figure 1. Compared to suburban tracts, central city zoning was relatively permissive, a strong majority of Chicago tracts (81%) receiving an index score of 4 or less. In contrast, suburban zoning was more restrictive, 89% of the suburban tracts receiving a score greater than 4.[6] Nonresi-

[4] Other analyses have demonstrated that specific regulations not accounted for in these methods of stratifying zoning categories tend to vary with the level of restrictiveness within each standard category. One may generally view any residential land-use category as a package of corresponding regulations where both the type of land use permitted and the minimum lot-area requirements set the limits on the absolute level of restrictiveness for each type of regulation (Shlay, 1981).

[5] We measured areas by the "dot grid method," which uses an overlay of uniformly distributed dots. We outlined the tract boundaries onto each zoning map, counted the number of dots within each zoning district within each tract, found the total number of dots within the tract, and formed a ratio of the number of dots in a particular zoning district to the total number of dots in the tract. For census tracts containing more than one local political jurisdiction, the scales of each zoning map containing the parts of the tracted area were used to convert the number of dots to a common scale of measurement.

[6] A score of "zero" is difficult to interpret within this scheme. Technically, zero indicates the most

Figure 1. The Distribution of Zoning Within Central City and Suburban Tracts in 1960



	Exclusionary Zoning Index			Percent Business[c]			Percent Manufacturing[d]		
	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N
Central City Tracts[b]	3.36	1.79	363	23.2	17.0	356	31.1	25.3	186
Suburban Tracts	6.18	1.62	146	12.8	14.1	139	16.0	20.5	78

[a] The index was computed to range from 0 to 11 where 11 is the most exclusive tract conceivable. Since the index is actually a continuous variable, the scores for each tract were rounded to the nearest whole number.
[b] These figures are based on a weighted sample of central city tracts. Low income–high density tracts and high income–low density tracts were undersampled in the city. By weighting these tracts by a factor of 1.78, these statistics describe a representative sample of central city tracted areas. The weighting procedure produces an artificially inflated sample size. The already representative suburban sample of tracts required no weighting.
[c] Means and standard deviations computed on the basis of only those tracts with some portion zoned for business land use.
[d] Means and standard deviations computed on the basis of only those tracts with some portion zoned for manufacturing land use.

dentially zoned land was also more available within the central city. Compared to suburban tracts, the average city tract that had any land zoned for either business or manufacturing land use allocated twice the proportion of land for these purposes.

These unsurprising city-suburb differences in the application of zoning provide one indicator that zoning has been effectively measured at the tract level. A second indicator is whether tract land-use regulations correspond to the actual land use contained within these areas. Assuming that zoning regulations are obeyed and enforced, the way in which land is zoned should be directly related to local housing density and composition.

Table 2 shows the regression of 1960 tract housing composition (number of housing units per square mile, percent single family dwellings and percent multiple family dwellings) on the three 1960 zoning variables. These "static" analyses do not represent causal processes but simply indicate the association between local zoning and land use. For each additional unit of the index, more exclusively zoned tracts were less dense (more than 2,000 fewer dwelling units per square mile in Chicago and more than 1,000 units in suburban areas) and contained proportionately more single-family homes than permissively residentially zoned tracts within both the central city and the suburbs. Although housing density did not tend to vary significantly according to the amount of land allocated for business use, tracts zoned heavily for business contained fewer dwelling units (90 in Chicago and 26 in the

permissively zoned tracts. A score of zero, however, also is obtained by tracts containing no land zoned for housing. While it can be argued that tracts containing only nonresidential zoning are, in fact, permissively zoned, the purpose of the exclusionary zoning index is to summarize the level of *residential* zoning exclusiveness within tracts. Therefore, for descriptive purposes, tracts receiving a score of zero are included in this discussion. In analytic calculations, these tracts are not included in analyses in which the Exclusionary Zoning Index is used.

Table 2. The Correspondence Between Zoning and Housing for Central City and Suburban Tracts in 1960

	No. of housing Units/Mile^2 1960		% Single-Family Dwellings 1960		% Multifamily Dwellings 1960[a]	
	City	Suburb	City	Suburb	City	Suburb
Independent Variables						
Exclusionary zoning index[b]	−2,020.04*	−1,022.96*	11.63*	12.27*	−13.34*	−6.64*
	(250.50)	(95.88)	(.64)	(1.17)	(.64)	(1.10)
% Zoned for business	42.60	−14.97	−.17*	−.29*	.28*	.23*
	(26.34)	(9.94)	(.07)	(.12)	(.07)	(.11)
% Zoned for manufacturing	−88.76*	−25.87*	−.04	−.10	−.007	−.03
	(16.69)	(7.15)	(.04)	(.09)	(.04)	(.08)
Constant	16,120.39*	9,502.02*	−9.76*	−4.22*	91.71*	58.14*
	(1,255.47)	(684.23)	(3.20)	(8.34)	(3.22)	(7.89)
R^2	.39	.51	.63	.57	.71	.33
N	237	144	237	144	237	144

NOTE: Entries are unstandardized regression coefficients with associated standard errors in parentheses.
[a] Number of units within structures containing three or more dwelling units.
[b] Index ranges from 1–11 where 1 is permissive residential zoning and 11 is exclusive residential zoning.
* Significant at .05 level.

suburbs) but more apartment houses (92 in Chicago and 58 in the suburbs) for each percentage increase in land zoned for manufacturing. The R^2 for each regression is high and significant, indicating that the derived zoning measures effectively capture local zoning practices at the census tract level.

Model Specification

The analyses which follow estimate the effects of tract zoning in 1960 on changes in tract housing and population composition over the decade. The general model employed is as follows:

$$T_{70_{ij}} = T_{60_{ij}} + \Sigma b_k C_{ik} + \Sigma b_m Z_{im} + e$$

where $T_{70_{ij}}$ and $T_{60_{ij}}$ are 1960 or 1970 measures j for tract i (e.g., number of housing units, median family income, number of nonwhites) and $\Sigma b_k C_{ik}$ is a vector of k tract characteristics for tract i that relate to tract change. $\Sigma b_m Z_{im}$ is a vector of m zoning variables for tract i in 1960, and e is the usual stochastic error term.

The purpose of the model is to capture the effects of the zoning regulations in force in 1960 on tract changes between 1960 and 1970. Since tracts tend to persist in character over time, do not change location, and have an existing physical plant, infrastructure, and the like, the 1960 characteristics of a tract are very good predictors of its 1970 characteristics. Indeed, the 1960 characteristics embody the effects of tracts' histories, namely the effects over time of the many urban social processes that determine the tract characteristics at a point in time. The model therefore states that the effects of zoning are to be measured by deviations from expected growth patterns based upon the persistence of characteristics over time (Bohrnstedt, 1969; Wright et al., 1979; Pendleton et al., 1980). When other characteristics of tracts are entered into the regressions such variables are to be regarded as exogenous to the 1960 characteristics. The coefficients for zoning measures derived from the model are estimates of the effects of zoning regulations net of the expected changes in the characteristics in question.

FINDINGS AND DISCUSSION

The "prosperous" sixties were a period of intense housing development and, in part, a period of housing "undevelopment" (Hartman, 1975; Mollenkopf, 1978). Urban renewal provided central cities with the legal mechanisms and federal funding for local reconstruction involving the demolition of large numbers of housing units. Rapid development of vacant suburban land was fostered by low inter-

est rates and improvements in intra-SMSA highway systems.

Newly built suburban housing and an increase in family real income meant that many urbanites were no longer constrained to live in the city. Thousands moved to new suburban housing developments, leaving behind older, less valuable housing situated in the seemingly unpleasant, dangerous, and dirty environment of central cities. In addition, newly formed households found that housing "bargains" were located mainly in the suburbs.

At the same time, another population movement was taking place, the migration of southern blacks into northern industrial urban centers. Although the number of blacks moving into the city was large, these new urbanites barely replaced the large number of middle class families moving out to the suburbs. The city gained a larger proportion of blacks but grew smaller in population. Table 3 summarizes these changes for the central city and suburban periphery from 1960 to 1970.

The Effects of Zoning on Local Housing Composition

Although the results of Table 2 indicate that the zoning classifications of tracts were reflected in the housing densities and housing stock compositions of those tracts in 1960, the issue we are concerned with is how those zoning regulations affected change in the period from 1960 to 1970. Do tracts with high exclusionary ratings become more or less dense in 1970 than one would expect on the basis of their 1960 densities? Are there corresponding changes in the expected housing stock composition? These issues are addressed in Tables 4 and 5.

Table 4 shows the effects of zoning on the 1970 densities of tracts. Suburban tracts become less dense than expected the higher the Exclusionary Zoning Index for that tract. Thus each additional level in average index values leads to about 253 fewer dwelling units per square mile in 1970. Each unstandardized regression coefficient represents the average net increase (or decrease) in the number of

Table 3. Changes in Selected Housing Population Characteristics for Central City and Suburban Tracts from 1960 to 1970

	City Tracts (N = 363)[a]					Suburban Tracts (N = 146)				
	1960	1970	Change 60–70			1960	1970	Change 60–70		
Selected Characteristic	$\bar{X}$	$\bar{X}$	$\bar{X}$	SD	R_{60-70}	$\bar{X}$	$\bar{X}$	$\bar{X}$	SD	R_{60-70}
Total housing units	1430	1444	+14	389	.93	1827	2277	+499	1010	.70
Housing density	9315	8975	−339	2911	.93	2685	2975	+290	640	.96
Number of single-family dwellings	324	332	+8	162	.95	1250	1387	+137	554	.87
Number of apartments (3 or more units)	824	846	+22	372	.92	349	625	+276	522	.60
Number of owner-occupied dwellings	456	485	+28	141	.97	1253	1525	+273	621	.84
Number of renter-occupied dwellings	916	961	+45	343	.93	575	835	+259	565	.60
Total population	4163	3987	−176	1225	.90	5922	6976	+1045	3174	.76
Population density	26614	24449	−2165	11021	.83	8044	8388	+345	1287	.97
Median family income	6550	9813	+3263	2337	.69	9119	13957	+4838	1708	.96
Number nonwhite	1059	1454	+395	1707	.73	158	300	+142	489	.83
Number single	741	861	+120	344	.84	880	1296	+416	610	.53
Number families with children under 18	564	385	−180	296	.75	864	892	+29	513	.76
Number families without children under 18	501	364	−139	271	.75	607	724	+118	320	.55

[a] These figures are based on a weighted sample of central city tracts. Low income–high density tracts and high income–low density tracts were undersampled in the city. By weighting these tracts by a factor of 1.78, these statistics describe a representative sample of central city tracted areas. The weighting procedure produces an artificially inflated sample size. The already representative suburban sample of tracts required no weighting.

Table 4. The Effects of 1960 Zoning on Changes in Housing Density from 1960 to 1970 for Central City and Suburban Tracts

	1970 Housing Density/ Units per Square Mile	
Independent Variables	City	Suburb
1960 Housing density (Housing units per mile²)	0.87*	0.95*
	(.03)	(.03)
Exclusionary zoning index[a]	−205.35	−253.01*
	(125.68)	(50.61)
% Zoned for business	−52.18*	−22.15*
	(11.75)	(3.93)
% Zoned for manufacturing	−26.58*	−4.18
	(7.84)	(2.93)
Constant	3262.09*	2325.74*
	(727.79)	(413.58)
R^2	.87	.94
N	237	144

NOTE: Shown are regression coefficients with their associated standard errors in parentheses.

[a] Index ranged from 1–11 where 1 is permissive residential zoning and 11 is exclusive residential zoning.

* Significant at .05 level.

housing units per square miles associated with a unit change in the corresponding independent variable. A "net decrease" does not mean that more exclusively zoned tracts experienced an absolute decrease in housing density, but that the increase in housing density was smaller in more exclusively zoned tracts compared to more permissively zoned tracts.

The coefficient for Chicago tracts is about the same size, but does not reach statistical significance. We suspect this shows that, with the scarcity of vacant land within Chicago tracts, the effects of zoning on growth are very variable. Hence outlying Chicago tracts reflect the same pattern as suburban tracts, and interior Chicago tracts are constrained by the lack of vacant land. Interior tracts that have experienced much demolition under urban renewal or other types of reconstruction, however, may be more responsive to the effects of zoning restrictions. The potential effects of both processes may be to produce more variability in response to zoning in Chicago tracts and hence larger standard errors.

The coefficients for business and manufacturing zoning also show that larger percentages of land zoned for those purposes in 1960 correlate with less growth in housing in 1970 than expected. The coefficients for business zoning are significant in both Chicago and the suburbs, but only in Chicago for manufacturing zoning. Each percent of business zoning meant 52 fewer housing units per square mile in Chicago and 22 fewer units per square mile in the suburbs in 1970, possibly reflecting the tendency for business establishments to drive out residential uses over time. These coefficients may represent the classical ecological process of business and manufacturing competing successfully for land use, but our framework suggests they are especially competitive when helped by appropriately permissive zoning.

Table 5 disaggregates these general effects by estimating separately the impacts of zoning on 1960–1970 changes in the number of single-family dwellings, the number of multiple-family dwellings (3 or more units per structure), the number of owner-occupied units and the number of renter-occupied units.

Included in each model are measures of tract areas (square miles) and measures of the changes in the total number of dwelling units contained within each tract over the decade. Tract area was included because the method of defining these units (by population size) tended to co-vary with tract zoning and because large tracts simply contain larger land areas available for potential development. Since census tract boundaries are defined in terms of the number of people living within these areas, large tracts (in area) tend to contain fewer people per square mile and hence have lower housing densities than small tracts. (The correlation between tract area and housing density in 1960 was −.40). Including tract area in the model thus enables the estimation of zoning effects that are not artifacts of the way in which census tracts were defined. Including the change in the total number of housing units over the decade accounts for general housing market forces so that the zoning effects are net of the tendency for some tracts to have experienced more growth than others. The zoning coefficients thus

Table 5. The Effects of Zoning on Housing Changes from 1960 to 1970 for Central City and Suburban Tracts

	Number of Housing Units 1970 for Selected Types of Housing							
	Single-Family Dwellings		Multifamily Dwellings[a]		Owner-Occupied Housing[b]		Renter-Occupied Housing	
	City	Suburb	City	Suburb	City	Suburb	City	Suburb
1960 Housing characteristic[c]	.78*	.92*	1.00*	1.06*	.92*	.93*	1.08*	1.10*
	(.04)	(.03)	(.01)	(.04)	(.03)	(.03)	(.01)	(.04)
Exclusionary zoning index[d]	41.93*	58.93*	−21.86*	−45.04*	15.09*	36.24*	−16.37*	−18.36
	(6.48)	(14.48)	(7.82)	(14.55)	(6.01)	(14.04)	(5.88)	(15.57)
% Zoned for business	.35	1.83	.81	−1.64	.02	.006	−.43	−.51
	(.57)	(1.32)	(.66)	(1.26)	(.54)	(1.34)	(.50)	(1.26)
% Zoned for manufacturing	−.90*	−.18*	−.37*	−.48	−.54	−.16	.26	−.51
	(.39)	(.95)	(.43)	(.91)	(.40)	(.96)	(.33)	(.90)
Change in total housing 1960–1970	.06*	.47*	.92*	.53*	.19*	.54*	.88*	.58*
	(.02)	(.02)	(.03)	(.02)	(.02)	(.02)	(.02)	(.02)
Tract area (square miles)	430.76*	18.56*	−177.95*	−15.84*	338.27*	21.75*	−226.60*	−17.50*
	(60.18)	(4.78)	(54.37)	(4.59)	(56.50)	(4.87)	(41.89)	(4.55)
Constant	−160.17*	−386.60*	150.71*	352.94*	−42.34	138.90	73.25*	95.20
	(26.67)	(91.54)	(39.03)	(108.33)	(24.68)	(93.76)	(30.53)	(122.13)
R^2	.93	.97	.98	.92	.98	.98	.99	.94
N	237	144	237	144	237	144	237	144

NOTE: Shown are regression coefficients with their associated standard errors in parentheses.
[a] Number of units within structures containing three or more dwelling units.
[b] Owner-occupied units include condominium units.
[c] The associated 1960 housing measures of each respective dependent variable.
[d] Index ranges from 1–11 where 1 is permissive residential zoning and 11 is exclusive residential zoning.
* Significant at .05 level.

represent the degree to which zoning influences the type of housing growth or the mix of housing within tract areas net of changes in the total housing stocks.

Within both the city and the suburbs, zoning directly influenced the type of housing development experienced from 1960 to 1970. More single family dwellings tended to be added to more exclusively zoned tracts, while more multifamily dwellings tended to be built within more permissively zoned tracts.

These effects are quite substantial and are independent of the housing-stock growth trends within tracts. Thus for each additional unit of the Exclusionary Zoning Index, a tract in Chicago experienced an additional 42 single-family homes, and a suburban tract about 59. In other words, two tracts of the same area that were 10 units apart on the index and grew at the same rate between 1960 and 1970 experienced in 1970 on the average a difference of between 420 and 590 single-family homes. Correspondingly large differences were also shown in multifamily dwellings, with coefficients of 22 fewer such units in Chicago and 45 in the suburbs for each higher level of the Exclusionary Zoning Index.[7]

Since tenure and dwelling types are closely related, the two regression equations involving owner-occupied and renter-occupied units reflect largely the same trends as noted above, although at reduced levels.[8] Indeed, since the Chicago SMSA during that period experienced a growth in condominiums (owner-occupied units in multiple-unit structures) this lack of direct correspondence is partly to be expected on that basis alone.

Neither central city nor suburban nonresidential zoning substantially influenced local housing development; these four regressions all have R^2 values in excess of .90. Since most of the variance among tracts in 1970 is explainable on the basis of tract variance in 1960, such high values are to be expected. Yet given the correlations shown in Table 4, it is also clear that zoning does contribute significantly to the explanation of tract status in 1970.

Zoning Effects On Population Composition Change and Growth

By regulating neighborhood housing composition, zoning indirectly regulates neighborhood socioeconomic composition. Higher prices for housing accompanying more restrictive land-use zoning can bar residence to lower income households. To the extent that race and socioeconomic level are related positively, restrictive zoning also eliminates blacks and other low-income racial and ethnic groups. Whether or not the passing of such zoning regulations was motivated by desires to segregate by income or by race or ethnicity is not a question we can answer with these data; all we can do is to estimate the consequences.

Table 6 models the effects of zoning on changes in median family income levels and changes in the number of nonwhite tract residents from 1960 to 1970. The change in the total population of tracts over the decade is included in the model of racial change in order to separate in-

---

[7] Analyses on changes in tracts' total housing stocks show that more exclusively zoned tracts tended to gain more housing units than permissively zoned tracts, although these coefficients are not statistically different from zero. Consistent with zoning's impact on changes in housing density, tracts zoned heavily for business and manufacturing gained fewer housing units compared to residentially zoned areas. Zoning's influences on changes in the total population of a tract tend to mirror zoning's impact on changes in the total number of housing units contained within these areas (Shlay, 1981).

[8] Other analyses examine zoning's influence on changes in the cost of tract housing over the decade, specifically changes in median rent levels and median owner-assessed housing values. Rent levels increased more within exclusively zoned tracts, which limited the amount of land available for rental housing. Rent levels increased less within permissively zoned tracts with little or no restrictions on high density development. Zoning's impact on housing value appreciation indicates that different processes were operating within Chicago and within the suburbs. Within suburban areas, housing values tended to appreciate more within exclusively zoned residential tracts. Within Chicago, housing values tended to appreciate more within permissively zoned residential tracts. These different effects may indicate that within an already "built up" central city where vacant land is scarce, land values tend to increase at a greater rate within areas zoned to permit competition for high-intensity land use (Shlay, 1981).

Table 6. The Effects of Zoning on Tract Socioeconomic Change from 1960 to 1970 for Central City and Suburban Tracts

	Median Family Income 1970		Number of Nonwhites 1970	
Independent Variables	City	Suburb	City	Suburb
1960 Socioeconomic characteristic	1.16*	1.33*	.95*	1.18*
	(.09)	(.04)	(.04)	(.07)
Exclusionary zoning index[a]	39.19	248.07*	−67.30	−37.50
	(110.83)	(83.35)	(57.40)	(33.78)
% Zoned for business	2.69	−7.17	4.09	−3.10
	(10.16)	(7.48)	(5.59)	(3.34)
% Zoned for manufacturing	−2.97	−1.18	−3.02	−.90
	(6.73)	(5.63)	(3.41)	(2.52)
Change in total population 1960–1970[b]	—	—	.88*	.02
	—	—	(.07)	(.02)
Tract area (square miles)	819.86	80.71*	1,547.46*	−4.93
	(812.31)	(23.82)	(437.47)	(12.14)
Constant	1,937.10*	177.12	417.68	379.57
	(709.11)	(539.57)	(277.98)	(235.20)
R^2	.51	.61	.72	.70
N	237	144	237	144

NOTE: Shown are regression coefficients with their associated standard errors in parentheses.
[a] Index ranges from 1–11 where 1 is permissive residential zoning and 11 is exclusive residential zoning.
[b] Change in total population not included in model of tract income-level changes.
* Significant at .05 level.

creases or decreases in nonwhites from overall population growth or decline.

Central city zoning did not directly influence 1960–1970 changes in average household income levels of tract residents. Suburban zoning, however, clearly worked to increase neighborhood income segregation. Restrictively zoned suburban tracts changed upwards in median income more than expected, but our data do not show whether this greater change reflects greater immigration of more affluent families or the greater upward income shifts of people living in such areas.

Neither central city nor suburban zoning worked to increase neighborhood racial segregation. We may note that the signs of the coefficients are negative, consistent with expectations that the more restrictively zoned neighborhoods had lower growths in nonwhite populations. There may be several reasons for this finding. First, the finding is consistent with racial segregation's being more a result of direct mechanisms than of zoning. For example, it may well be that real estate agents and brokers simply bar the way to nonwhites by their referral policies. Secondly, segregation may result from the search patterns employed by blacks, in which they seek only in neighborhoods close by to black neighborhoods—a pattern which reflects sectoral migration rather than wider dispersal among many neighborhoods with equivalent housing offerings. Indeed, as Table 6 indicates, most of the variance in the number of nonwhites in 1970 is taken up with the 1960 version of that variable, reflecting the fact that blacks move where there are blacks.

The Effect of Zoning on Life-Cycle Composition

Since housing requirements change as households go through typical life cycles, zoning may affect life cycle–related aspects of population composition by restricting or enhancing appropriate housing types. Thus a neighborhood that has mainly small, one-bedroom apartments is scarcely regarded as a desirable location for couples with several school-age children. Similarly, zoning that only permits single-family homes on large lots may attract to a residential area couples with many school-age children.

Census tract statistics allow only indirect indicators of life-cycle stages, as in the three indicators of tract life-cycle composition in Table 7: changes in the number of single adults, changes in the number of childless families, and changes

Table 7. The Effects of Zoning on Family Life-Cycle Compositional Changes with Tracts from 1960 to 1970 for Central City and Suburban Tracts

	Number of Single Persons 1970		Number of Families Without Children 1970		Number of Families with Children Less Than 18 Yrs. 1970	
Independent Variables	City	Suburb	City	Suburb	City	Suburb
1960 Population characteristic[a]	1.13*	1.11*	.56*	1.12*	.46*	1.21*
	(.03)	(.05)	(.03)	(.11)	(.03)	(.08)
Exclusionary zoning index[b]	−26.47*	7.00	15.40	56.85*	23.44*	19.90
	(8.28)	(11.75)	(8.66)	(19.89)	(7.38)	(25.79)
% Zoned for business	−3.07*	−1.41	.18	1.09	−.36	2.62
	(.81)	(1.16)	(.88)	(1.94)	(.74)	(2.58)
% Zoned for manufacturing	−1.04*	1.01	−1.44*	−1.06	−1.22*	−1.86
	(.53)	(.85)	(.57)	(1.38)	(.46)	(1.84)
Change in total population 1960–1970[c]	.36*	.25*	.06*	.04*	.03*	.08*
	(.02)	(.01)	(.01)	(.01)	(.01)	(.01)
Tract area (square miles)	−80.28	−2.36	239.03*	5.95	303.96*	26.98
	(68.37)	(3.97)	(68.48)	(6.91)	(58.07)	(9.44)
Constant	244.94*	42.49	10.28	−370.01*	14.94	−426.57*
	(43.93)	(98.00)	(46.24)	(175.65)	(37.60)	(187.33)
R^2	.92	.95	.68	.51	.69	.77
N	237	144	237	144	237	144

NOTE: Shown are regression coefficients with their associated standard errors in parentheses.
[a] The associated 1960 population measure of each respective dependent variable.
[b] Index ranges from 1–11 where 1 is permissive residential zoning and 11 is exclusive residential zoning.
[c] The examination of the change in the number of single persons accounts for the change in the total number of adults (14 years and older) from 1960 to 1970.
* Significant at .05 level.

in the number of families with children. The first equation includes changes in tract adult population to hold constant overall increases in adults. The remaining models include changes in total population. Again, these population-change characteristics account for population trends per se, so that the effects of zoning are over and above general population trends.

Central city zoning worked to distribute urban residents according to stages in the family life cycle. The number of single persons tended to increase more in permissively zoned residential areas, while the numbers of married people with or without children tended to increase more in restrictively zoned neighborhoods.[9]

Suburban zoning's influence on change in tract life-cycle composition was less pronounced. Tracts gained neither single persons nor families with children due to local zoning. Rather, exclusively zoned suburban neighborhoods tended to have larger increases in numbers of families who had no children in the household.

These city-suburb differences emphasize the difference between urban and suburban residential environments. Permissively zoned central city areas, accompanied by traffic, noise, and an absence of open space, may have been regarded as especially unsuited for raising young children. By contrast, the more exclusively zoned suburbs were less dense and did not offer the externalities offered by central city neighborhoods, so that suburban families were less restricted in their choice of a suitable residential environment for rearing children.[10] The at-

[9] Whether these differential changes occur through population movements into or out of the tracts cannot be ascertained. Probably both processes are occurring at the same time; unattached adults moving into permissively zoned areas and families with children moving out.

[10] The differences between central city and suburban residential environments are suggested by the distribution of residential zoning within tract areas. Within the central city, tracts zoned for residential land use tended to be either entirely permissively zoned or entirely restrictively zoned. Within the suburbs, however, few tracts were entirely permissively zoned. Residentially zoned suburban tracts tended to be dominated by zoning for single-family dwellings and contained small amounts of land for high-density land use. Hence, suburban tracts with low Exclusionary Index Scores were apt to contain some single family housing (Shlay, 1981).

traction of childless, suburban families to more exclusively zoned neighborhoods may reflect that these couples (professional working couples or older couples in the empty-nest stage) simply had more money to spend on housing than growing families.

CONCLUSIONS AND IMPLICATIONS

According to Sussna (1966), "We had better recognize that the humble zoning ordinance is probably of direct concern to more people than any other statute." Our examination of zoning provides good reasons for such concern. Zoning is an important, local, political method for shaping the housing and population composition of metropolitan neighborhoods. An overview of the findings presented illuminates the way in which this process works.

(1) By limiting housing density, zoning directly influences increases in the availability of rental housing.

(2) Zoning methods create increasingly segregated neighborhoods in terms of housing tenure. Restrictively zoned neighborhoods tend to become increasingly neighborhoods of homeowners, while less restrictively zoned neighborhoods become more intensively neighborhoods of renters.

(3) Central city zoning does not tend to increase neighborhood income segregation. Urban zoning simply increases trends in which households are sorted according to stages in the family life cycle.

(4) Suburban zoning tends to increase neighborhood income segregation.

(5) Residential zoning does not increase racial segregation either in the central city or in the suburbs. Although zoning does not directly influence neighborhood racial segregation, suburban zoning may maintain already established patterns by including or excluding the types of people who can afford to live within differently regulated areas.

These findings are a mixture of the surprising and the quite expectable. On the one hand, some of the effects of zoning are as intended by the advocates of zoning. Exclusionary zoning practices tend to "keep up the neighborhood" by segregating by housing types, tenure, and nonresidential uses. On the other hand, there are some unexpected results; socioeconomic and racial effects are either nonexistent or not uniform. Especially surprising was the finding that nonwhites increase in neighborhoods not any more than expected on the basis of 1960 composition, regardless of the exclusionary zoning practices in place.

Our analyses also show that the political processes that result in zoning modify market effects. While our analyses do not directly address the question of how great these zoning effects are compared, say, to locational factors that get expressed in housing and land prices, these effects are considerable in affecting the growth of housing stocks in neighborhoods. Local political decision making can modify market forces, at least as far as trends over a decade's time are concerned.

APPENDIX

Municipalities and Counties Containing Tracted Areas

City
Chicago

Peripheral Municipalities
Arlington Heights
Aurora
Bedford Park
Berwyn
Burr Ridge
Calumet City
Chicago Heights
Cicero
Des Plaines
Downers Grove
Elgin
Elmhurst
Evanston
Evergreen Park
Forest Park
Glencoe
Glen Ellyn
Glenview
Highland Park
Hometown
Joliet
LaGrange Park
LaGrange
Lansing
Lyons
Maywood
Melrose Park
Morton Grove
Naperville
Niles
Oak Park
Park Ridge
River Forest
River Grove
Skokie
Steger
Stickney
Waukegan
Wilmette

Counties
Cook County
DuPage County

REFERENCES

Babcock, Richard
1969 The Zoning Game. Madison, Wisconsin: University of Wisconsin Press.

Bassett, Edward M.
1936 Zoning: The Laws, Administration, and

Court Decisions During the First Twenty Years. New York: Russell Sage Foundation.
Berry, Brian J. L. and John D. Kasarda
1977 Contemporary Urban Ecology. New York: MacMillan.
Bergman, Edward
1974 Eliminating Exclusionary Zoning: Reconciling Workplace and Residence in Suburban Areas. Cambridge, MA: Ballinger Publishing Company.
Beuscher, Jacob H., Robert R. Wright, and Morton Gitelman
1976 Cases and Materials on Land Use. St. Paul: West Publishing Company.
Bohrnstedt, George W.
1969 "Observations on the measurement of change." Pp. 113–29 in Edgar F. Borgatta (ed.), Sociological Methodology. San Francisco: Jossey-Bass, Inc.
Danielson, Michael B.
1976 The Politics of Exclusion. New York: Columbia University Press.
Davidoff, Paul and Linda Davidoff
1971 "Opening the suburbs: toward inclusionary land use controls." Syracuse Law Review 22:509–35.
Delafons, John
1969 Land Use Controls in the United States. Cambridge, MA: The MIT Press.
Downs, Anthony
1970 Urban Problems and Prospects. Chicago: Markham.
1973 Opening Up the Suburbs. New Haven: Yale University Press.
Fitch, Robert
1977 "Planning New York." Pp. 246–84 in Roger E. Alcaly and David Mermelstein (eds.), The Fiscal Crisis of American Cities. New York: Vintage Books.
Form, William H.
1954 "The place of social structure in the determination of land use." Social Forces 32:317–23.
Friedman, Lawrence
1968 Government and Slum Housing. Chicago: Rand McNally.
Haar, Charles
1959 Land Use Planning: A Casebook on the Use, Misuse and Reuse of Urban Land. Boston: Little, Brown.
Hartman, Chester W.
1975 Housing and Social Policy. Englewood Cliffs, N.J.: Prentice Hall.
Harvey, David
1973 Social Justice and the City. Baltimore: Johns Hopkins Press.
Hawley, Amos
1950 Human Ecology. New York: Ronald Press.
Logan, John R.
1976 "Industrialization and the stratification of cities in suburban regions." American Journal of Sociology 82:333–48.
1978 "Growth, politics and the stratification of places." American Journal of Sociology 84:404–16.
1980 "Growth and succession in suburban communities." The Sociological Quarterly 21:93–105.
Makielski, S. J.
1966 The Politics of Zoning: The New York Experience. New York: Columbia University Press.
Mollenkopf, John
1978 "The postwar politics of urban development." Pp. 117–52 in William Tabb and Larry Sawers (eds.), Marxism and the Metropolis: New Perspectives in Urban Political Economy. New York: Oxford University Press.
Molotch, Harvey
1976 "The city as a growth machine: toward a political economy of place." Americal Journal of Sociology 82:309–32.
Moskowitz, David H.
1977 Exclusionary Zoning Litigation. Cambridge, MA: Ballinger.
National Commission on Urban Problems
1968 Building the American City. Washington, D.C.: Government Printing Office.
Nelson, Robert H.
1977 Zoning and Property Rights. Cambridge, MA: Massachusetts Institute of Technology Press.
Neutze, Max
1968 The Suburban Apartment Boom—Case Study of a Land Use Problem. Baltimore: The Johns Hopkins Press.
Park, Robert, Ernest W. Burgess, and Roderick D. McKenzie
1967 The City. Chicago: The University of Chicago Press.
Pendleton, Brian F., Richard D. Warren, and H. C. Chang
1980 "Correlated denominators in multiple regression and change analysis." Pp. 80–103 in Edgar F. Borgatta and David J. Jackson (eds.), Aggregate Data Analysis and Interpretation. Beverly Hills: Sage Publications.
Perin, Constance
1977 Everything in Its Place: Social Order and Land Use in America. Princeton, N.J.: Princeton University Press.
Schafer, Robert
1975 "Exclusionary land use controls: conceptual and empirical problems in measuring the invisible wall." Dept. of City and Regional Planning, Harvard University, Discussion Paper D-75-2.
Schwirian, Kent P.
1974a "Some recent trends and methodological problems in urban ecological research." Pp. 7–18 in K. P. Schwirian, Comparative Urban Structure. Lexington, MA: D. C. Heath.
1974b Comparative Urban Structure. Lexington, MA: D. C. Heath.
Shlay, Anne B.
1981 "Zoning for whom: the impact of zoning on tract housing and population changes from 1960 to 1970 within the Chicago SMSA." Ph.D. Dissertation. University of Massachusetts/Amherst.

Sussna, Stephen
1966 "Fifty years of zoning." American Bar Association Journal 152:1030–3.
Toll, Seymour
1969 Zoned America. New York: Grossman Publishers.
Willhelm, Sidney
1962 Urban Zoning and Land Use Theory. New York: The Free Press.
Williams, Norman
1966 The Structure of Urban Zoning and Its Dynamics in Urban Planning and Development. New York: Buttenheim.
1975 American Land Planning Law. Chicago: Callaghan and Company.

Williams, Norman, Jr. and Thomas Norman
1971 "Exclusionary land use controls: the case of northeastern New Jersey." Syracuse Law Review 22:475–507.

Wright, James D., Peter H. Rossi, Sonia R. Wright, and Eleanor Weber-Burdin
1979 After the Clean-Up: Long-Range Effects of Natural Disasters. Beverly Hills: Sage Publications.

STRATIFICATION, WORK, AND VALUES: A POLISH–UNITED STATES COMPARISON[1]

KAZIMIERZ M. SLOMCZYNSKI
University of Warsaw and *National Institute of Mental Health*

JOANNE MILLER
MELVIN L. KOHN
National Institute of Mental Health

In Poland and the United States social stratification is related to parental values and to social orientations, with men of higher position more likely to value self-direction and to have a social orientation consonant with valuing self-direction: a nonauthoritarian perspective, personally responsible standards of morality, and trustfulness. These relationships result in large measure from the greater opportunities afforded by higher position to be self-directed in one's work.

In the United States, higher social-stratification position is associated with more favorable self-conceptions, largely as a result of the greater opportunities for occupational self-direction that higher position affords. In Poland, lower position is associated with greater self-confidence and less anxiety.

Polish sociologists have studied the social-stratification system of their socialist society intensively, focusing on the extent of social inequality and the magnitude of correlation among various dimensions of social standing. (For comprehensive reviews of this literature in English see Telenbach, 1974; Wesolowski and Slomczynski, 1977; and Slomczynski and Krauze, 1978.) Although some attention has been given to the psychological concomitants of social inequality (e.g., Firkowska, et al., 1978) and to the psychological aspects of social structure (e.g., Nowak, 1969), there has, until now, been no research addressing the relationships of social stratification and job conditions to general values, self-conceptions, and social orientations. For this reason Wesolowski (1975:98) proposed a replication in Poland of Kohn and Schooler's United States study of social stratification and personality (Kohn, 1969; Kohn and Schooler, 1969) to see whether their findings apply as well to socialist society. This proposed replication was conducted in 1978.

In their original United States study, Kohn and Schooler (1969; Kohn, 1969: Chapters 4, 5, 9, and 10) found that men of higher "social-class" position are more likely than are men of lower "social-class" position to value self-direction and to have self-conceptions and social orientations predicated on the belief that self-direction is both possible and efficacious. They also found that the relationships of social-stratification position to values, self-conceptions, and social orientations result from the cumulative impact of the idea-

* Direct all correspondence to Slomczynski-Miller-Kohn, Bldg. 31, Room 4-C-11, N.I.H., Bethesda, MD 20205.

We are indebted to Wlodzimierz Wesolowski, who initiated the Polish replication and encouraged the investigators throughout the inquiry; to Krystyna Janicka and Jadwiga Koralewicz-Zebik, who coauthored with Slomczynski the research proposal, with Krystyna Janicka subsequently taking major responsibility for administering the study; to Pawel Wejland and Pawel Danilowicz, who conducted the pilot study; and to Andrzej Mokrzyszewski, Anna Sianko, and Stanislawa Walkowska, who took responsibility for fieldwork, coding, and preparation of the data. We are also indebted to Ronald Schoenberg, for advice on causal modelling; to Cheryl Keller, Carrie Schoenbach, Bruce Roberts, Margaret Renfors, and Diane Mueller, for computer programming and statistical analysis; to Virginia Marbley, for transcribing innumerable drafts of this paper; and to Carmi Schooler, Carrie Schoenbach, Karen A. Miller, Jeylan T. Mortimer, and Ronald Schoenberg, for critical readings of an earlier version of this manuscript. The models in this paper were estimated by MILS, an advanced version of LISREL (Joreskog and van Thillo, 1972), developed by Ronald Schoenberg.

tional flexibility and breadth of perspective afforded by educational experience and the opportunities for self-direction in one's work afforded by higher occupational position. Moreover, the relationships of stratification to psychological functioning were substantially attributable to three job conditions that affect the exercise of self-direction in one's work—namely, the substantive complexity of the work, the closeness of supervision, and the routinization of the work. This they interpreted as resulting from a learning-generalization process by which the lessons of the job are generalized to men's views of themselves and of the larger society. Experiencing self-direction in one's work leads to valuing self-direction, off as well as on the job; to seeing society as so constituted that responsible individual action is practicable; and to seeing oneself as competent, effective, and in control of the forces that affect one's life.

Kohn (1969:196) hypothesized that occupational self-direction would play a major part in explaining the relationship of social stratification to values and orientation in all "sizeable industrial societies." Subsequent studies carried out in the United States (Mortimer and Lorence, 1979a,b; Miller, et al., 1979) and in other capitalist countries (see Grabb, 1981; Hynes, 1977; Hoff and Grueneisen, 1977; Bertram, 1976a,b; Coburn and Edwards, 1976; and the review in Kohn, 1981) have confirmed the Kohn-Schooler findings and supported Kohn's hypothesis. But there has until now been no test of the generalizability of the findings to socialist society. Does social stratification bear the same relationship to valuation of self-direction in socialist society as in capitalist society, or might the ethos of individuality be less related to stratification position in a society whose economic system is not predicated on the alleged virtues of individual enterprise? Even if the relationship of stratification to values and orientation is the same in socialist as in capitalist society, is the explanation of this relationship necessarily the same? Or might it be, for example, that, while opportunities for self-direction in one's work are of pivotal explanatory importance in a capitalist economic system, in a socialist system other mechanisms intervene—mechanisms specific to national ownership of the means of production under central planning?

METHODS OF THE UNITED STATES AND POLISH SURVEYS

The original United States survey was based on interviews conducted in 1964 with a representative sample of 3101 men employed in civilian occupations throughout the country (see Kohn, 1969, for details). These data provide the primary U.S. base for the comparative Polish–U.S. analyses. In addition, for analyses of the reciprocal effects of occupational self-direction and psychological functioning, we utilize a ten-year follow-up study conducted in the United States with a representative subsample of 687 of the men originally interviewed in 1964 (Kohn and Schooler, 1978).

The Polish survey was conducted in 1978 under the auspices and with the financial support of the Polish Academy of Sciences. It was designed to be an exact replication of the main parts of the U.S. study. Questions pertaining to parental values, self-conceptions, and social orientations as well as to occupational self-direction were directly adopted from the Kohn-Schooler 1964 questionnaire. The measures of social stratification—formal education, job income, and occupational status (prestige)—came from previous Polish studies (see Danilowicz and Sztabinski, 1977; Slomczynski and Kacprowicz, 1979) where they had been intensively tested.

The initial translation of the American questionnaire into Polish involved a thorough assessment of the meaning of entire questions and of particular phrases. This assessment was especially valuable for translating questions that include colloquial expressions, such as "going to pieces," "end up causing trouble," and "feel upset," which have several equivalents in Polish. Several alternative translations of each question were judged collectively by a group of linguistic experts, who evaluated the semantic and syntactic equivalence of the Polish and English versions. The resulting version of

the questionnaire was subjected to a pilot study, based on interviews with fifty persons selected from the upper and lower ends of the educational and occupational distributions. The questionnaire was then modified and again pre-tested, this time on university students, and then reviewed by experienced interviewers in group discussion.

The Polish sample was designed to represent men, aged 19–65, living in urban areas, and employed full-time in civilian occupations. Although the rural peasantry is not represented, farmers living in proximity to urban centers are included, making the Polish sample more comparable to the U.S. sample than a sample fully representative of Poland. A final sample of 1557 men was obtained through a multistage probability sampling scheme.

INDEX CONSTRUCTION: PARENTAL VALUES AND ORIENTATION TO SELF AND SOCIETY

Parental Values

By parental values, we mean those standards of desirability that parents would most like to see embodied in their children's behavior (see Kohn, 1969:18–20). There is every reason to believe (see Kohn, 1969: Chapter 4; Kohn, 1977) that people's values for their children reflect their more general values, not only for children, but also for themselves.

To assess parental values, both the U.S. and Polish surveys asked fathers to select from a standard list of 13 characteristics those that they consider to be most desirable and those they consider to be least desirable for a boy or girl the same age as a randomly selected child of their own (for the rationale of this mode of inquiry, see Kohn, 1969:Chapter 4). They were asked to partially rank the 13 characteristics by selecting the three most desirable, the one of those three that is most desirable of all, the three least important, and the one of those three that is least important of all. This information made it possible to score each characteristic from 5 for most valued to 1 for least valued, with the mid-rank signifying that the respondent did not select the characteristic either as desirable or as not desirable. Extensive pretests verified that the 13 characteristics used in the American survey are appropriate to Polish fathers. The set of characteristics is given in Table 1.

In constructing measurement models of parental values for the U.S. and Polish data, we begin with exploratory factor analysis, from which we learn that there are two main substantive factors in both bodies of data—valuation of self-direction versus conformity to external authority and valuation of striving for success. Since the U.S. data contain also a nonsubstantive age or "maturity" factor, we control age by partialling child's age out of the variance-covariance matrices on which the confirmatory factor analyses are based.

The forced-choice nature of the parental-values inquiry results in linear dependency—one necessarily knows how a father will rate any characteristic if one knows how he rates the remaining set. To avoid linear dependency, we must leave some characteristics out of the variance-covariance matrix on which the model is based, at minimum one pair of characteristics that are highly correlated with each other. (See Kohn and Schoenbach, 1980; Jackson and Alwin, 1980; and Alwin and Jackson, 1981). For both countries, we drop a pair whose components connote the two ends of the self-direction/conformity factor—self-control and being a good student for the United States, and responsibility and neatness for Poland. To reduce the problem of linear dependency even further, without distorting the factor structure, we exclude from the U.S. model three characteristics that do not relate to either factor in the exploratory factor analysis—honesty, acting like a boy (or girl) should, and ability to get along well with other children. Since, for Poland, all characteristics are related to one or the other factor, we exclude two characteristics—manners and consideration—that Kohn and Schoenbach's (1980) multinational analysis of parental values indicates may suffer in translation.

The measurement models, summarized in Table 1, confirm the existence of both

Table 1. Measurement Model of Parental Values for American and Polish Fathers (Standardized coefficients, all paths statistically significant, p ≤ .05)

	Standardized Path: Concept to Indicator	
	American Fathers (N = 1499[a])	Polish Fathers (N = 660[a])
Self-direction versus Conformity to External Authority		
That (s)he has good manners [That (s)he is socialized in proper way (colloquial expression implying well brought up child.)]	−.48	—
That (s)he has good sense and sound judgment [That (s)he has sound judgment of (her) himself and others.]	.25	.37
That (s)he obeys parents well	−.33	−.39
That (s)he is responsible	.41	—
That (s)he is considerate of others [That (s)he takes others into consideration.]	.42	—
That (s)he is interested in how and why things happen [That (s)he tries to understand things precisely.]	.26	.19
That (s)he is neat and clean	−.40	—
That (s)he is a good student	—	−.45
That (s)he acts like a girl (boy) should	—	−.16
That (s)he has self-control	—	.27
Striving for Success		
That (s)he tries hard to succeed	1.00	1.00
That (s)he is responsible	−.19	—
That (s)he is considerate of others [That (s)he takes others into consideration.]	−.24	—
That (s)he is interested in how and why things happen [That (s)he tries to understand things precisely.]	−.14	—
That (s)he is honest	—	−.19
That (s)he gets along well with other children	—	−.14
Chi-square (Degrees of freedom)[b]	62.84(9)	173.39(24)
Ratio chi-square per d.f.	6.98	7.22

NOTE: Question wording in Polish interview appears in brackets when a modification of American wording was made.

[a] Questions on parental values were asked only of parents with a child 3–15 years of age and refer to a randomly selected child in that age range.

[b] Correlations between residuals not shown.

self-direction/conformity and striving-for-success dimensions in parental values in the two countries, which is consistent with data from a sample of Warsaw mothers and also with data from several other countries (Kohn and Schoenbach, 1980). The measurement models also provide Bartlett-type factor weights (Lawley and Maxwell, 1971:109–11) from which we construct indices for use in the causal analyses.[1]

[1] All causal analyses reported in this paper are based on variance-covariance matrices. The covariances based on Bartlett-type factor scores were corrected for attenuation, using estimates of the correlations of factor scores with true scores provided by the measurement models.

Orientations to Self and Society

In both the U.S. and Polish surveys, the indices of orientation to self and society are based on a battery of 57 questions, mainly of the "agree-disagree" and "how often?" types, with multiple response categories. In the original U.S. analyses, Kohn and Schooler developed indices of nine dimensions of orientation (Kohn, 1969, Appendix D; or Kohn and Schooler, 1969). These dimensions are authoritarian conservatism, standards of morality, trustfulness, receptivitity to change, idea conformity, self-confidence, self-deprecation, anxiety, and fatalism.

In the present analysis, we initially use orthogonal, exploratory factor analysis to verify the existence of the same factors in

both the U.S. and Polish data and to search for the best indicators of each concept in each country. Thus, for example, in exploratory factor analyses that extract as few as seven to as many as fourteen factors, authoritarian conservatism consistently emerges as a factor in both sets of data, with many of the same items correlating substantially with this factor in both countries. Some items, however, appear to be more strongly related to authoritarian conservatism in the U.S. data and others appear to be more strongly related to authoritarian conservatism in the Polish data (see J. Miller et al., 1981). Similarly, for other dimensions the core items are much the same in both countries, but the strength of relationship of individual items to any particular dimension differs.

On the basis of the exploratory factor analyses, we have identified subsets of items that best measure several of the dimensions of orientation in the Polish data. Each subset has then been used in a confirmatory factor analysis (Joreskog, 1969) to produce a measurement model of the specific construct. This procedure has enabled us to produce measurement models for five concepts: authoritarian conservatism, anxiety, self-confidence, self-deprecation, and idea-conformity (see Table 2). For two other concepts, standards of morality and trustfulness, there is no clear factor present in the orthogonal, exploratory factor analyses of the Polish data, but a priori selections of items yield satisfactory measurement models of these concepts. Thus, we have indexed all but two of the nine facets of orientation to self and society of the original U.S. study.

All of the Polish measurement models

Table 2. Measurement Models of Social Orientation and Self-conception for American and Polish Men (Standardized coefficients, all paths statistically significant, $p \leq .05$)

	Standardized Path: Concept to Indicator			
	American Men (1964) (N = 3101)		Polish Men (1978) (N = 1557)	
Concepts and Indicators[a]	U.S. Model	Polish Model	Polish Model	U.S. Model
Authoritarian Conservatism				
The most important thing to teach children is absolute obedience to their parents.	.65	.71	.73	.66
In this complicated world, the only way to know what to do is to rely on leaders and experts. (M)	.52	.49	.53	.51
It's wrong to do things differently from the way our forefathers did. (M)	.44	.40	.40	.52
Any good leader should be strict with people under him in order to gain their respect.	.46	.52	.53	.54
No decent man can respect a woman who has had sex relations before marriage.	.37	.34	.42	.45
Prison is too good for sex criminals; they should be publicly whipped or worse. (M)	.39	—	—	.41
Young people should not be allowed to read books that are likely to confuse them.	.44	—	—	.33
There are two kinds of people in the world: the weak and the strong.	.62	—	—	.41
People who question the old and accepted ways of doing things usually just end up causing trouble. (M)	.51	—	—	.21
One should always show respect to those in authority.	—	.31	.62	—
You should obey your superiors whether or not you think they're right.	—	.33	.50	—
Do you believe that it's all right to do whatever the law allows, or are there some things that are wrong even if they are legal? (M)	—	.17	.42	—
Correlation: U.S. model/Polish model	.88		.90	
Chi-square (Degrees of freedom)[b]	11.6 (17)	15.5 (15)	9.35 (13)	21.3 (18)
Ratio chi-square per d.f.	.68	1.03	.72	1.18

Table 2. Continued

	Standardized Path: Concept to Indicator			
	American Men (1964) (N = 3101)		Polish Men (1978) (N = 1557)	
Concepts and Indicators[a]	U.S. Model	Polish Model	Polish Model	U.S. Model
Personally Responsible Standards of Morality				
It's all right to do anything you want as long as you stay out of trouble.	−.62	−.60	−.72	−.61
If something works, it doesn't matter whether it's right or wrong.	−.44	−.44	−.37	−.44
It's all right to get around the law as long as you don't actually break it.	−.60	−.62	−.28	−.32
Do you believe that it's all right to do whatever the law allows, or are there some things that are wrong even if they are legal? (M)	−.36	—	—	−.31
Correlation: U.S. model/Polish model.	.96		.93	
Chi-square (Degrees of freedom).	0.0 (1)	d.f. = 0	d.f. = 0	4.4 (1)
Ratio chi-square per d.f.	.0	d.f. = 0	d.f. = 0	4.40
Trustfulness				
Do you think that most people can be trusted?	.75	.49	.31	—[c]
If you don't watch out, people will take advantage of you.	−.31	−.48	−.33	—[c]
Human nature is really cooperative. (M)	.15	.15	.31	—[c]
When you get right down to it, no one cares much what happens to you. (M)	—	−.35	−.35	—[c]
Correlation: U.S. model/Polish model	.84		—	
Chi-square (Degrees of freedom)	d.f. = 0	11.8 (2)	.99 (1)	—[c]
Ratio chi-square per d.f.	d.f. = 0	5.92	.99	—[c]
Idea Conformity[d]				
According to your general impression, how often do your ideas and opinions about important matters differ from those of your relatives? (M)	.50		.58	
How often do your ideas and opinions differ from those of your friends? (M)	.76		.55	
How about from those of other people with your religious background?	.57		.41	
Those of most people in the country? (M)	.40		.63	
Correlation: U.S. model/Polish model.	—		—	
Chi-square (Degrees of freedom).	d.f. = 0		d.f. = 0	
Ratio Chi-square per d.f.	d.f. = 0		d.f. = 0	
Self-esteem, Two-Concept Model				
Self-confidence				
I take a positive attitude toward myself. (M)	.55	.41	.39	.38
I feel that I'm a person of worth, at least on an equal plane with others.	.50	.68	.58	.57
I am able to do most things as well as other people can.	.36	.50	.53	.55
I generally have confidence that when I make plans I will be able to carry them out.	.58	—	—	.33
Once I have made up my mind, I seldom change it.	—	.14	.41	—
Correlation: U.S. model/Polish model	.82		.93	
Self-deprecation				
I feel useless at times.	.46	.52	.53	.54
At times I think I am no good at all.	.57	.80	.71	.70
There are very few things about which I'm absolutely certain.	.39	.26	.29	.29
I wish I could be as happy as others seem to be.	.45	—	—	.12
I wish I could have more respect for myself. (M)	.60	—	—	.26
Correlation: U.S. model/Polish model	.81		.99	
Chi-square (Degrees of freedom)[e]	16.8 (15)	4.8 (4)	4.8 (8)	6.1 (11)
Ratio chi-square per d.f.	1.12	1.20	0.60	0.55

Table 2. Continued

Concepts and Indicators[a]	American Men (1964) (N = 3101) U.S. Model	American Men (1964) (N = 3101) Polish Model	Polish Men (1978) (N = 1557) Polish Model	Polish Men (1978) (N = 1557) U.S. Model
	Standardized Path: Concept to Indicator			
Anxiety				
How often do you feel that you are about to go to pieces? (M)	.58	.61	.53	.51
How often do you feel downcast and dejected? (M)	.73	.54	.60	.61
How frequently do you find yourself anxious and worrying about something? (M)	.56	.56	.54	.57
How often do you feel uneasy about something without knowing why?	.55	.54	.58	.57
How often do you feel so restless that you cannot sit still?	.47	.50	.48	.49
How often do you find that you can't get rid of some thought or idea that keeps running through your mind? (M)	.40	.52	.47	.48
How often do you feel bored with everything?	.52	.48	.41	.40
How often do you feel powerless to get what you want out of life? (M)	.48	.45	.42	.41
How often do you feel guilty for having done something wrong?	.34	—	—	.14
How often do you feel that the world just isn't very understandable?	.46	—	—	.46
How often do you feel that there isn't much purpose to being alive?	.50	—	—	.37
Correlation: U.S. model/Polish model.	.92		.98	
Chi-square (Degrees of freedom).	14.5 (19)	12.3 (12)	15.8 (13)	31.3 (28)
Ratio chi-square per d.f.	0.76	1.03	1.22	1.12

[a] A high score on the indicator generally implies agreement or frequent occurrence; where alternatives are posed, the first alternative is scored high. "M" denotes a slight modification of American question wording in the Polish interview.

[b] Correlations between residuals not shown.

[c] It was not possible to solve the "U.S. Model" with the Polish data.

[d] The U.S. and Polish models of idea conformity have the same indicators.

[e] The chi-square of the 2-concept model of self-esteem refers to the entire model.

fit the data well by the chi-squared goodness-of-fit test. We nevertheless have greater confidence in the indices based on the more rigorous selection procedure—those for authoritarian conservatism, anxiety, self-confidence, self-deprecation, and idea-conformity—and less confidence in the measurement models of standards of morality and trustfulness. Not surprisingly, the indices that meet our more rigorous criteria tend to be those for which there are a substantial number of indicators.

The same procedure applied to the U.S. data—developing a measurement model of each concept, using items suggested by exploratory factor analysis—yields similar results (see Table 2).

To evaluate whether the Polish and U.S. measurement models are indexing the same concepts, wherever possible we have developed two sets of alternative measurement models, imposing the model suggested by the exploratory factor analysis of each country on the data of the other country. It was not possible to impose the U.S. model of trustfulness on the Polish data. In every other instance, the imposed model fits the data well, albeit not quite as well as does the model derived from exploratory factor analysis of the country's own data. Moreover, the correlations between indices based on the two alternative operationalizations are generally high, especially for the indices based on larger numbers of overlapping indicators (again see Table 2). The evidence is thus considerable that the con-

cepts measured in the Polish replication are essentially the same as those measured in the American study.

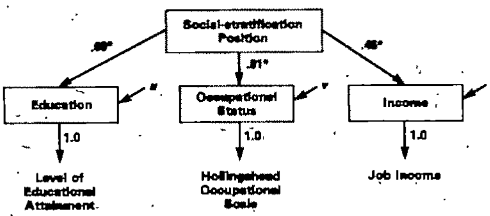
We think of these seven facets of orientation as falling along a continuum, from those that are focused preponderantly on society (for example, authoritarian conservatism) to those that are focused preponderantly on self (for example, self-confidence). Most of the seven partake of both, with some (notably, idea-conformity) being more or less equally concerned with society and with self. Even though the seven facets of orientation do fall along a continuum, it is useful in discussing them to follow Kohn and Schooler's practice of dichotomizing them into one group termed social orientations (authoritarian conservatism, personally responsible standards of morality, and trustfulness) and another group termed orientations to self or self-conception (idea-conformity, self-confidence, self-deprecation, and anxiety). "Social orientation" and "self-conception" are merely rubrics. In particular, self-conception is meant to be broader than "self-esteem" and even broader than "self-concept," as those terms are employed by Rosenberg (1979) and other students of the self.

SOCIAL-STRATIFICATION POSITION

Measurement Models

In this analysis of social stratification, we modify both the terminology and the indices that Kohn and Schooler originally employed. Referring to "aggregates of individuals who occupy broadly similar positions in a hierarchy of power, privilege, and prestige," Kohn and Schooler (1969:660) used the term "social classes." We reserve that term to mean a group defined in terms of its relation to the means of production (see Wesolowski, 1969 and 1979; and Kohn, 1977 or 1981). The hierarchical distribution of power, privilege, and prestige is more properly called social stratification.

For both the Polish and the U.S. data, we have developed measurement models in which social-stratification position is treated as a "second-order" concept, the first-order concepts being education, occupational status, and income (see Figures 1 and 2). The U.S. model uses only one indicator of each first-order concept: the level of educational attainment, the Hollingshead occupational scale, and income from the main job. The Polish model uses a comparable index of education; for occupational status, though, there are two indicators—the Polish Prestige Score (Slomczynski and Kacprowicz, 1979) and the International Prestige Score (Treiman, 1977); and for income, the indicators are earnings from the main job and total job income.[2]



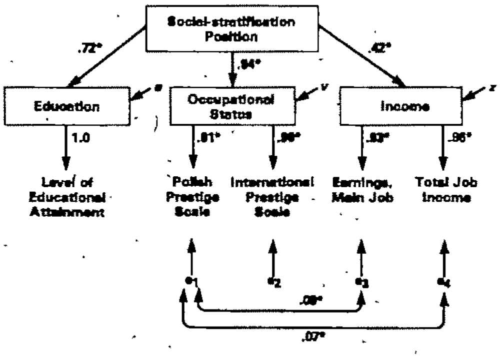
*Statistically significant, $p \leq .05$.
χ^2 (d.f.) = — (0)

Figure 1. Measurement Model of Social-Stratification Position for U.S. Men (N = 3101, standardized coefficients)

Despite the U.S. model having only single indicators of occupational status and income, while the Polish model has two indicators of each, the models are essentially comparable. The Polish model

[2] The correlation of Treiman's International Prestige Scale with a prestige scale developed for Poland is lower than with prestige scales developed for specific capitalist countries (Treiman, 1977:176). Our model nevertheless suggests that Treiman's Scale is as valid a measure of occupational status in Poland as is the Polish Prestige Scale (Slomczynski and Kacprowicz, 1979). Treiman's Scale measures a more universal pattern of occupational prestige than does the Polish Scale, which reflects specific changes within post–World War II Polish society (see Sarapata and Wesolowski, 1961; Pohoski et al., 1976).

The use of the Hollingshead index of occupational status for research in the U.S. is validated by longitudinal measurement models that show the Hollingshead index to be as strong an indicator of occupational status as is Treiman's International Prestige Scale, the Hodge-Siegel Index, or the Duncan Socioeconomic Index (Schooler, 1980; Kohn and Schoenbach, 1980).



*Statistically significant, $p \leq .05$.
χ^2 (d.f.) = 5.8(1)

Figure 2. Measurement Model of Social-Stratification Position for Polish Men (N = 1557, standardized coefficients)

shows that standardized paths from first-order concepts to their indicators are exceedingly strong (ranging from .91 to .96), clearly implying that one-indicator models would do almost as well in this rare situation where every indicator is highly effective. Of crucial importance in comparing the Polish and U.S. models, the pattern of relationships of social-stratification position with education, occupational status, and income is nearly the same in the two countries. We conclude that the basic social phenomenon being measured is much the same in Poland and the United States and that these models provide entirely comparable indices of social-stratification position in the two countries.

Relationship to Values, Self-conceptions, and Social Orientations

The relationships of social-stratification position to parental values and to social orientations are remarkably similar in Poland and the United States (see Table 3). For example, the correlation between social-stratification position and father's valuation of self-direction for their children is .46 in both countries, with education accounting for the lion's share of these correlations. In both countries, higher position is associated with valuing self-direction, in contrast to valuing conformity to external authority. The correlation between social-stratification position and parental valuation of striving for success is close to zero in both countries. It is parents' valuation of self-direction, not of striving for success, that is tied to their social-stratification position (see the discussions in Kohn, 1977 and 1981).

The correlations of social-stratification position with authoritarian conservatism, standards of morality, and trustfulness are very similar in the two countries; holding higher position is consistently associated with having self-directed orientations to

Table 3. Relationships Between Social-Stratification Position and Values and Orientations for Polish (N = 1557) and American (N = 3101) Men (except for parental values, where Ns are 660 and 1499)

	Standardized Beta Coefficients in Multiple-Regression Equations						Zero-Order Correlation Using Second-Order Construct	
	Education		Occupational Status		Income			
	Poland	U.S.	Poland	U.S.	Poland	U.S.	Poland	U.S.
Values								
Parental valuation of self-direction	.39*	.30*	.10*	.13*	.05*	.16*	.46*	.46*
Social Orientation								
Authoritarian conservatism	−.41*	−.48*	−.13*	−.11*	−.03	−.07*	−.50*	−.54*
Standards of morality	.21*	.27*	.25*	.14*	−.01	.09*	.44*	.42*
Trustfulness	.13*	.19*	.27*	.22*	.08*	.06*	.42*	.42*
Self-conception								
Idea conformity	−.16*	−.18*	.04	.00	−.07*	−.02	−.13*	−.15*
Self-confidence	−.06	.09*	−.07*	.09*	.00	.09*	−.12*	.22*
Self-deprecation	−.00	−.16*	−.10*	−.11*	−.05	−.07*	−.12*	−.29*
Anxiety	.14*	.06*	−.04	−.13*	.01	−.04*	.08*	−.12*

* Statistically significant, $p \leq .05$.

the larger society. The relationships between social-stratification position and self-confidence (r = +.22), but in Poland different in the two countries. In the United States, men of higher position are more likely to have favorable self-conceptions, just as they are more likely to have favorable orientations to the larger society. Even in the United States, however, the magnitudes of the correlations with social stratification are notably smaller for self-conceptions than for social orientations. In Poland, not only are the magnitudes of the correlations with stratification small, but unlike the United States, higher position is not uniformly associated with more favorable self-conceptions. The most notable cross-national difference is in the relationship between social-stratification position and self-confidence. In the United States, higher position is associated with more self-confidence (r = +.22), but in Poland the opposite is true (r = −.12). Similarly, albeit less strikingly, for anxiety: In the United States, higher position is associated with somewhat less anxiety (r = −.12), but in Poland higher social position is associated with slightly greater anxiety (r = +.08). The difference between the two countries is less pronounced for self-deprecation, higher position being associated with less self-deprecation in both countries. Still, the relationship is stronger in the United States (r = −.29) than in Poland (r = −.12). There is essentially no difference between the United States and Poland with respect to idea-conformity: In both countries, higher position is associated, to a modest but significant degree, with greater independence in one's ideas. But, then again, idea-conformity partakes as much of social orientation as of self-conception.

We believe that the differences between Poland and the United States in the relationships between social-stratification position and self-conception cast no doubt on the comparability of indices, because the U.S. and Polish measurement models of these concepts are very similar. Rather, we believe there may be differences between the United States and Poland in how social conditions facilitate or inhibit a favorable self-conception, particularly in how they affect self-confidence. We shall return repeatedly to this issue.

THE RELATIONSHIP BETWEEN SOCIAL STRATIFICATION AND OCCUPATIONAL SELF-DIRECTION

Measurement Models of Occupational Self-direction

At the heart of the original analysis of the U.S. data was Kohn's (1963) hypothesis that the relationships between social class and values and orientation are substantially attributable to those class-associated conditions of occupational life that facilitate or deter the exercise of initiative, thought, and independent judgment in one's work—that is, occupational self-direction. We think of occupational self-direction as primarily dependent upon the substantive complexity of the work, how closely it is supervised, and its degree of routinization. Since we see occupational self-direction as an overarching concept encompassing these three job conditions, we have developed multi-indicator measurement models for the United States and Poland that incorporate all three (see Table 4).

Substantive complexity of work is the degree to which performance of the work requires thought and independent judgment. Substantively complex work by its very nature requires making many decisions that must take into account ill-defined or apparently conflicting contingencies. For both the United States and Poland, our information about the substantive complexity of work is based on detailed questioning of each respondent about his work with things, with data or ideas, and with people (see Kohn, 1969:153–5, 271–6; or Kohn and Schooler, 1978). These questions provide the basis for seven ratings of each man's job: three appraisals of the complexity of the work in his job—with things, with data, and with people; an evaluation of the overall complexity of his work, regardless of whether he works primarily with data, with people, or with things; and an estimate of the amount of time he spends working at each of the three types of activity. The seven ratings are treated in Table 4 as indicators of the underlying but

Table 4. Measurement Models of Occupational Self-Direction for American and Polish Men (Standardized coefficients, all paths statistically significant, $p \leq .05$)

	Standardized Path: Concept to Indicator	
	American Men, 1964[b] (N = 3101)	Polish Men, 1978 (N = 1557)
First-Order Concepts and Indicators[a]		
Substantive Complexity		
Hours data	.45	.59
Hours things	−.59	−.64
Hours people	.43	.29
Complexity data	.82	.87
Complexity things	.23	.25
Complexity people	.78	.89
Overall complexity	.85	.85
Routinization		
Variability (predictability) of tasks	1.00	1.00
Closeness of supervision		
Free to disagree with supervisor	−.46	−.26
Self-rated closeness of supervision	.65	.55
Supervisor tells what to do	.60	.58
Importance of doing what told	.65	.40
Second-Order Concept and Indicators		
Occupational self-direction		
Substantive complexity	.98	.86
Routinization	−.24	−.39
Closeness of supervision	−.69	−.73
Chi-square (Degrees of freedom)	191.4 (53)	77.9 (34)
Ratio chi-square per d.f.	3.61	2.29

[a] Correlations between residuals not shown.
[b] U.S. model includes information about the substantive complexity of earlier jobs, which is not shown.

not directly measured construct, the substantive complexity of that job.

Closeness of supervision limits one's opportunities for occupational self-direction: Workers cannot exercise occupational self-direction if they are closely supervised, although not being closely supervised does not necessarily mean that they are required or even free to use initiative, thought, and independent judgment. In both studies, closeness of supervision is measured by a worker's subjective appraisals of his freedom to disagree with his supervisor, how closely he is supervised, the extent to which his supervisor tells him what to do rather than discussing it with him, and the importance in his job of doing what one is told to do.

Routinization is the final facet of occupational self-direction; highly routinized (repetitive and predictable) jobs restrict possibilities for exercising initiative, thought, and judgment, while jobs with a variety of unpredictable tasks may facilitate or even require the use of self-direction. We use slightly different measures of routinization for the United States and Poland. For the United States, respondents' work was coded from most variable (the work involves doing different things in different ways and one cannot predict what may come up) to least variable (the work is unvaryingly repetitive). For Poland, we do not include predictability in the index.

The measurement models for occupational self-direction (including error-correlations not shown in Table 4) provide a good fit to the variance-covariance matrix of the indicators in both countries. The relationships between first-order concepts and their indicators (e.g., the substantive complexity of work and the number of hours spent working with data, or closeness of supervision and the respondent's appraisal of whether his supervisor "tells him what to do") are quite similar for the two countries. So, too, are the relationships between the second-order concept, occupational self-direction, and the first-order concepts—substantive complexity, closeness of supervision, and routinization. In both countries, substantive complexity is the most powerful indicator of occupational self-direction, and routinization is the least powerful. There is a decided similarity in the structuring of work in the two countries, despite the obvious and important contrast in ownership of economic enterprise.

Reconceptualizing the Relationship Between Social Stratification and Occupational Self-direction

Kohn and Schooler (1969) asked whether substantive complexity, closeness of

supervision, and routinization might be mediating factors that explain—at least in substantial part—the relationships between social-stratification position and values and orientation. Their way of answering this question was to statistically control these three job conditions, to see how much the correlations between social-stratification position and values and orientation would thereby be reduced. Such an analytic procedure, however, assumes unidirectional causality, from social-stratification position to occupational self-direction. Although in the United States this may be an appropriate assumption for education, it may not be in Poland; and in both countries occupational status and job income not only affect but may also be affected by occupational self-direction. On the one hand, because the status and income derived from a job signify that job's placement in the organization of work and the system of social stratification, they affect the actual conditions of work experienced in the job. On the other hand, status and income are also rewards, distributed in relationship to critical aspects of the work performed, in particular, its substantive complexity, closeness of supervision, and routinization; hence, status and income can be affected by job conditions.[3]

We have modelled the interrelationship of current occupational self-direction and current social-stratification position, as depicted in Figure 3. In this model, current occupational self-direction is a second-order concept, its "indicators" being substantive complexity, closeness of supervision, and routinization. Current social-stratification position is now disaggregated into education and occupational position, the latter a second-order concept with status and income as "indicators." In Poland, it is quite common for men to continue their formal education well into their careers; accordingly, the model includes both the respondent's education at the beginning of his work career and his education at the time he started his current job. The initial level of education is treated as exogenous, but the educational level at the time the respondent began his current job is endogenous, because it can have been affected by previous job experience. Since most U.S. men complete their formal education before getting very far into their work careers, for the United States, only the level of education as of the time of the interview is employed. It is treated as exogenous to both current occupational position and current occupational self-direction.[4]

To assess the contemporaneous reciprocal effects of current occupational position and current occupational self-direction, we make use of somewhat limited, retrospective information about the occupational position of the job held by the the respondent immediately preceding the one held at the time of the interview and the degree of occupational self-direction he experienced in that job. The model permits the occupational position of the immediately preceding job to directly affect the present job's occupational position, but not its occupational self-direction; correspondingly, the occupational self-direction of the immediately preceding job is allowed to directly affect the present job's degree of occupational self-direction but not its occupational position. Allowing the "stabilities" but not the "cross-lagged" effects provides the identification needed to assess the reciprocal effects of current occupational position and current occupational self-direction. (For a discussion of the rational and also the dangers of these assumptions, see Heise, 1975: 184–5). The contemporaneous relationship between occupational position and occupational self-direction is undoubtedly the outgrowth of longer-term processes of selective recruitment into and retention in particular jobs, individual career mobility, and changes in the

[3] Even though our measures of occupational status are based on ratings for the occupation as a whole, we are treating them as proxies for job status. Job income is of course specific to the particular job. Thus, status and income are treated as attributes of the job; it is reasonable to think of them as subject to the contemporaneous effects of other job conditions.

[4] Our longitudinal data show a correlation of .97 between men's educational levels at the time of the first interview in 1964 and at the time of the follow-up interview ten years later. This correlation is essentially the same for the oldest, intermediate, and youngest segments of the work force.

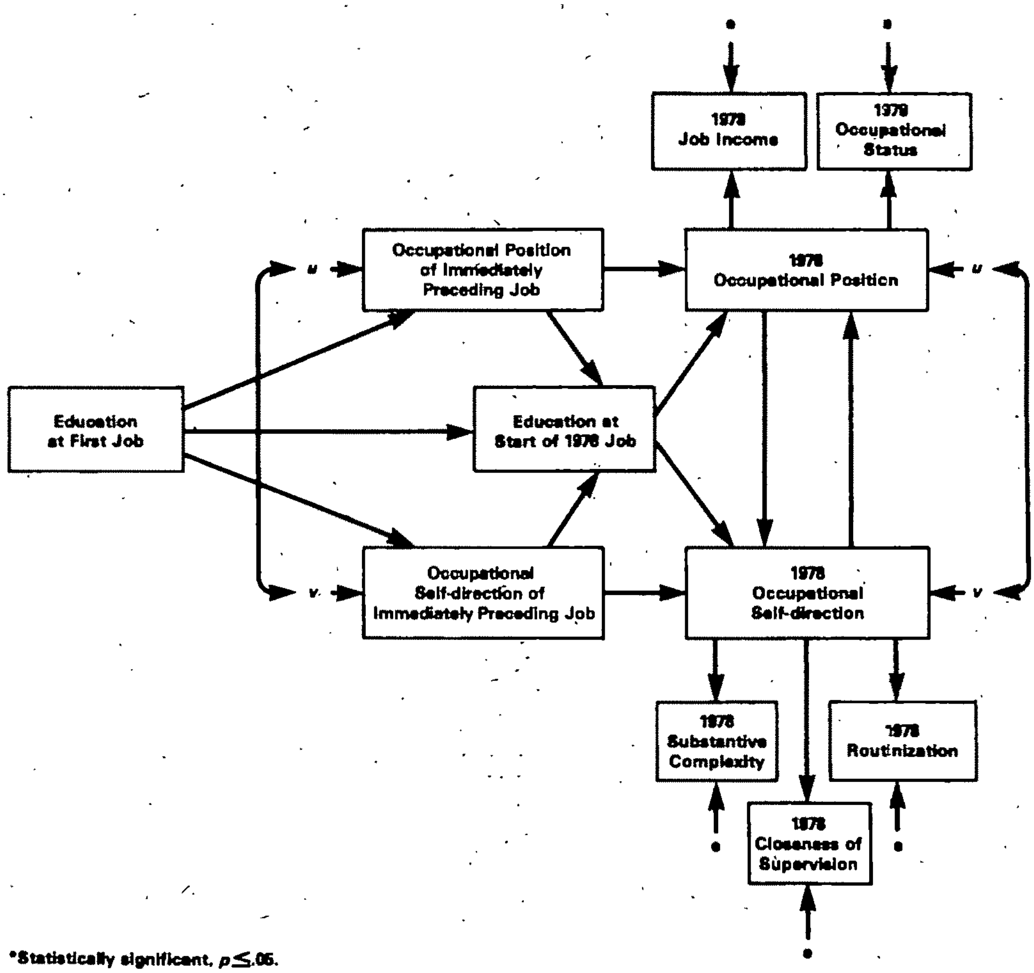
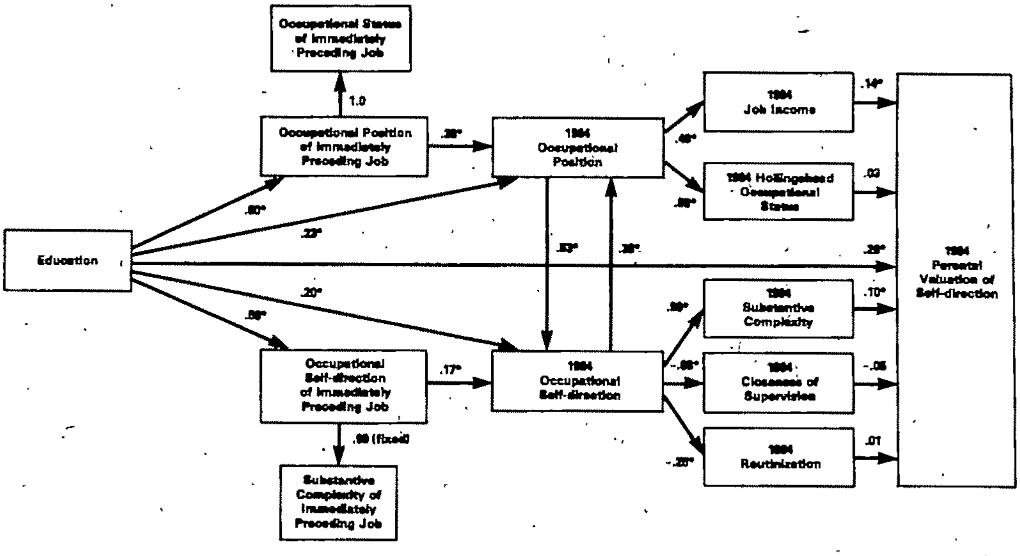


Figure 3. Diagram Illustrating Interrelationship of Occupational Self-direction and Social-Stratification Position

structuring of work. Our model depicts the contemporaneous outcomes of these ongoing processes.

We also introduce into the model those social characteristics that might be pertinent to job placement at this stage of career. They are allowed to affect both current occupational position and current occupational self-direction. For the United States, we consider age, race, national background, and religious background. For Poland, only age is pertinent. We include these social characteristics primarily as statistical controls.

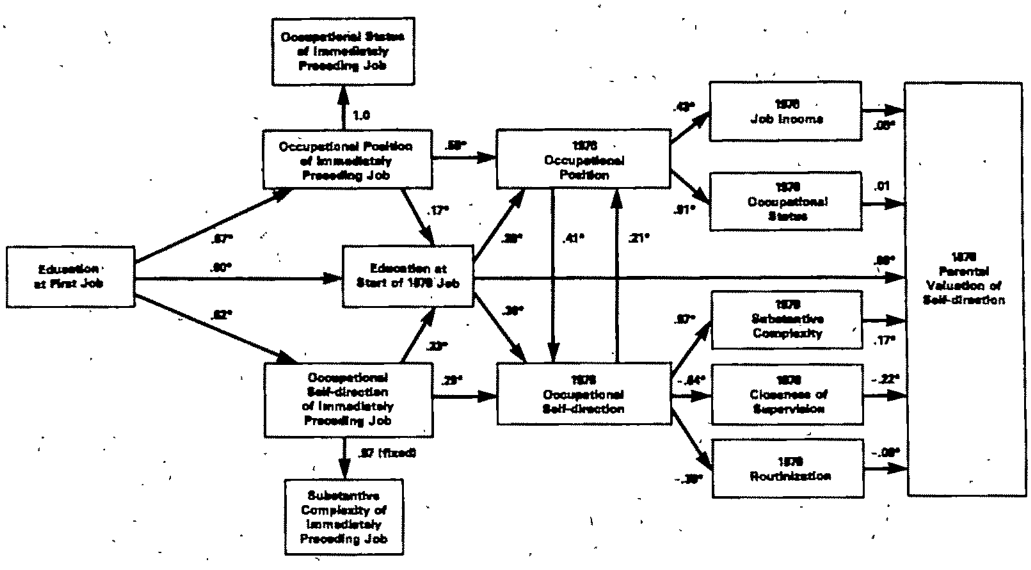
We find that, in both countries, much of education's effect on occupational self-direction occurs early in the career, in helping determine early job placement (see Figure 4 for the U.S. and Figure 5 for Poland). The path from education to the occupational self-direction of the immediately prior job is of approximately similar magnitude in the United States (.56) and Poland (.62). But then the processes diverge. In the United States, where for most men the formal educational process stops before the occupational career gets substantially under way, education continues to affect occupational self-direction, both directly and indirectly through occupational position, but is not itself affected by occupational self-direction. In Poland, where many men continue their formal educations during their occupational careers, educational attainment is part of an ongoing process, in which the educational level attained before the work career begins affects the



*Statistically significant, $p \leq .05$.

Figure 4. Effects of Education, Occupational Position, and Occupational Self-direction on Parental Valuation of Self-direction, for U.S. Men (N = 1499, standardized coefficients. Paths from background characteristics to endogenous variables not shown. Correlations between residuals not shown.)

occupational self-direction experienced early in the career, which in turn affects further education, which in turn affects the occupational self-direction of the current job. As a result, the educational level at the time of beginning the current job has a stronger direct effect on current occupational self-direction in Poland than does educational level in the United States. Nevertheless, the long-term effect of education on occupational self-direction is substantially the same in both countries:



*Statistically significant, $p \leq .05$.

Figure 5. Effects of Education, Occupational Position, and Occupational Self-direction on Parental Valuation of Self-direction, for Polish Men (N = 660, standardized coefficients. Paths from background characteristics to endogenous variables not shown. Correlations between residuals not shown.)

The cumulative effects of initial level of education on current occupational self-direction are .73 for Poland and .67 for the United States.[5]

Consonant with current education having a somewhat greater direct effect on occupational self-direction in Poland than in the United States, occupational position has a somewhat smaller direct effect in Poland (.41 as compared to .53). This difference, such as it is, may result in part from our possibly underestimating the stability of occupational self-direction in the United States and overestimating it in Poland (a function of the nature of the retrospective information available to us). But even if the difference is real, the core finding is that—despite our statistically controlling education, prior occupational self-direction, and relevant social characteristics—the effect of occupational position on occupational self-direction is considerable in both countries. Moreover, the reciprocal effect of occupational self-direction on occupational position is also substantial (.38 in the United States and .21 in Poland.) The result of these mutually reinforcing processes is that the estimated correlation between the occupational position and the occupational self-direction of the current job is extraordinarily high in both countries: .97 in the United States and .99 in Poland. This confirms a central tenet of Kohn's interpretation, that occupational self-direction is closely linked to social-stratification position, regardless of political or economic system.

EFFECTS OF SOCIAL-STRATIFICATION POSITION AND OCCUPATIONAL SELF-DIRECTION ON VALUES AND ORIENTATION

Separating the effects of occupational position and occupational self-direction on values, self-conceptions, and social orientations is a matter of considerable theoretical importance. Do the prestige and income of men's jobs affect their values and their orientations, with occupational self-direction mattering only incidentally? Or is it the reverse, that job conditions directly affect values and orientations, with prestige and income mattering not for their own sake but because the job's position in the social-stratification order markedly influences the worker's opportunities to be self-directed in his work? Even a near-unity correlation between occupational position and occupational self-direction in both countries does not necessarily mean that their high correlation results from the same processes or that the causal dynamics of occupational position and occupational self-direction vis-a-vis other variables are the same.

The method we use to differentiate the direct effects of occupational position from those of occupational self-direction is to disaggregate both phenomena into their components and allow these components to affect values, self-conceptions, and social orientations. Figures 4 and 5 provide examples that show the effects of education, occupational position, and occupational self-direction on fathers' valuation of self-direction for children. Here, again, we allow social characteristics that might be pertinent to current job placement to affect occupational position and occupational self-direction. These, as well as other social characteristics that might affect psychological functioning, are also allowed to affect values, self-conceptions, and social orientations.[6]

Consider first the model for U.S. men.

[5] Calculating direct, indirect, and total effects in nonrecursive models is more difficult than in recursive models because of feedback effects. The MILS computer program calculates total and indirect effects, following Fox (1980).

[6] Specifically, for the United States the pertinent social characteristics are age, race, national background, religious background, parental family socioeconomic status (mother's and father's education, father's occupational status, maternal and paternal grandfathers' occupational statuses), the urbanness and region of the country of the principal place where the respondent was raised, and the number of brothers and sisters he had. Some of the social characteristics in the U.S. model are omitted from the Polish model because they are irrelevant in the more culturally homogeneous Polish society (namely, race, religious background, and national background) and others are omitted because information is unavailable (mother's education, grandfathers' occupational statuses, region of origin, and number of children in the family of orientation).

Education and job income have sizeable and statistically significant direct effects (.26 and .14) on parental valuation of self-direction, while substantive complexity has a smaller but still significant effect (.10). Thus, education, one component of occupational position, and one component of occupational self-direction significantly affect parental valuation of self-direction, independent of the others.

In the model for Poland, only current education—not education as of the first job—is modelled as directly affecting parental valuation of self-direction. Thus, the modelling of direct effects on parental values is exactly comparable to that for the United States, even though the earlier part of the model is different. As in the model for U.S. fathers, education, income, and substantive complexity have statistically significant effects on fathers' valuation of self-direction; in Poland, so too do closeness of supervision and routinization. In Poland, as in the United States, education, occupational position, and occupational self-direction all have significant direct effects on parental valuation of self-direction.

Similar models for the three facets of social orientation (see Table 5) yield consistent results: In both countries, at least one of the components of social-stratification position significantly affects every aspect of social orientation. The components of occupational self-direction likewise have many statistically significant effects. In particular, with only a single exception, substantive complexity

Table 5. Effects of Social-Stratification Position and Occupational Self-Direction on Values and Orientations, Controlling Background Characteristics for Polish and American Men

	Standardized Path Coefficients					
	Stratification Position			Occupational Self-Direction		
	Education	Occ. Status	Income	Substantive Complexity	Closeness of Supervision	Routinization
Values						
Parental valuation of self-direction						
Poland	.09*	.01	.05*	.17*	−.22*	−.06*
U.S.	.26*	.03	.14*	.10*	−.05	.01
Social Orientation						
Authoritarian conservatism						
Poland	−.19*	.05	−.03	−.21*	.21*	.04
U.S.	−.32*	−.00	−.04*	−.22*	−.01	.03
Standards of morality						
Poland	.21*	.09	−.05	.08	−.11*	−.06*
U.S.	.30*	−.06	.04*	.12*	−.13*	−.05*
Trustfulness						
Poland	−.04	.18*	.07*	.12*	−.11*	−.12*
U.S.	.12*	.07*	.01	.11*	−.04	−.03
Self-conception						
Idea conformity						
Poland	−.07	.09	−.07*	−.09	.08*	−.01
U.S.	−.09*	.09*	−.01	−.17*	.03	.01
Self-confidence						
Poland	−.08	−.01	.00	−.07	−.07*	.03
U.S.	.00	−.02	.06*	.11*	−.18*	−.01
Self-deprecation						
Poland	−.04	−.12*	−.05	.11	.05	.01
U.S.	−.07*	.00	−.03	−.11*	.13*	.04*
Anxiety						
Poland	.14*	−.06	−.00	−.01	−.03	−.06*
U.S.	.02	−.15*	−.04	.09*	.07*	.02

NOTE: For Poland, background characteristics are age, urbanness of place raised, and father's occupational status. For the U.S., background characteristics are age, urbanness and region of place raised, religious background, race, national background, father's occupational status, maternal and paternal grandfathers' occupational statuses, and number of siblings.

* Statistically significant, $p \leq .05$.

significantly affects all three facets of social orientation in both countries. Closeness of supervision and routinization often add to that effect.

The picture, once again, is different for self-conceptions. In neither country is the pattern as clearcut for self-conceptions as for social orientations. For the most part, though, in the United States greater opportunity for occupational self-direction is associated with more favorable self-conceptions, even with social-stratification position statistically controlled. In Poland, there is little indication of occupational self-direction having much independent effect on self-conceptions. Substantive complexity does not have a statistically significant effect on any of the four facets of self-conception. Closeness of supervision, though, does increase idea-conformity and decrease self-confidence. Routinization affects anxiety, but by decreasing, not increasing it. One could hardly conclude that occupational self-direction favorably affects self-conceptions in Poland, independent of stratification position.

It would appear, then, that the answer to our question, does social-stratification position affect values and orientations independent of, or because of, its close association with occupational self-direction is: both. Education and occupational position affect values and social orientations even when occupational self-direction is statistically controlled. Thus, the psychological effects of social-stratification position cannot be entirely explained in terms of the greater opportunities for occupational self-direction available to people of higher educational and occupational position. But occupational self-direction does have an independent effect on values and social orientations. Some part of the effect of education and of occupational position on values and social orientations must therefore be indirect, a result of education and occupational position affecting occupational self-direction and of occupational self-direction, in turn, affecting values and social orientations. In the United States, this must be true for self-conceptions, too. The question becomes, then, what proportions of the effects of education and of occupational position are indirect, via their causal impact on occupational self-direction and its causal impact on values, self-conceptions, and social orientations?

By reaggregating the processes depicted in Figures 4 and 5 and Table 5, we can assess the degree to which the effects of education and of occupational position are mediated by occupational self-direction. To do this, we calculate the total effects of education and of occupational position on values and orientation, the indirect effects via occupational self-direction, and, from this information, the proportions of the total effects that are attributable to occupational self-direction. As shown in Table 6, in both the United States and Poland the effects of occupational position on parental values and on social orientations are substantially attributable to occupational self-direction. Concretely, in the United States, occupational self-direction explains somewhat more than half the effect of occupational position on parental valuation of self-direction; in Poland the corresponding proportion is four-fifths. Similarly for social orientations: In both countries, at least two-fifths and in some instances virtually the entire effect of occupational position on the particular facet of social orientation is attributable to occupational self-direction. The situation, again, is different for self-conceptions. In the United States, occupational self-direction largely explains the effects of occupational position on self-confidence and self-deprecation. It also explains the effects of occupational position on idea-conformity, but here the total effect is trivial. It is irrelevant to anxiety. In Poland, the total effects of occupational position on idea-conformity, self-confidence, and anxiety are little more than zero. In the one case where the total effect of occupational position is non-trivial—self-deprecation—occupational self-direction is irrelevant. Thus, in the United States, occupational self-direction largely explains the effects of occupational position on two pivotal facets of self-conception: self-confidence and self-deprecation; in Poland, occupational self-direction has no such explanatory power.

Our assessment of the role of occu-

Table 6. Effects of Education and of Occupational Position Attributable to Occupational Self-direction

	U.S. Men							
	Effects of Occupational Position				Effects of Education			
	Total Effect	Direct Effect	Indirect Effect	Indirect as %	Total Effect	Direct Effect	Indirect Effect	Indirect as %
Values								
Parental valuation of self-direction	.20	.09	.11	54	.41	.26	.11	27
Social Orientation								
Authoritarian conservatism	−.17	−.02	−.15	86	−.48	−.32	−.15	31
Standards of morality	.11	−.03	.14	100	.43	.30	.14	34
Trustfulness	.18	.07	.11	61	.26	.12	.11	42
Self-conception								
Idea conformity	−.03	.08	−.11	100	−.16	−.09	−.11	66
Self-confidence	.17	.01	.16	95	.17	.00	.16	98
Self-deprecation	−.16	−.01	−.14	92	−.22	−.07	−.15	65
Anxiety	−.15	−.15	−.01	6	−.06	.02	−.01	15
	Polish Men							
	Effects of Occupational Position				Effects of Current Education			
	Total Effect	Direct Effect	Indirect Effect	Indirect as %	Total Effect	Direct Effect	Indirect Effect	Indirect as %
Values								
Parental valuation of self-direction	.19	.03	.15	82	.24	.09	.14	60
Social Orientation								
Authoritarian conservatism	−.12	.04	−.16	100	−.33	−.19	−.15	45
Standards of morality	.15	.06	.08	57	.30	.21	.08	26
Trustfulness	.31	.19	.12	39	.11	−.04	.11	100
Self-conception								
Idea conformity	−.00	.05	−.05	100	−.11	−.07	−.05	46
Self-confidence	−.03	−.01	−.02	57	−.10	−.08	−.01	14
Self-deprecation	−.11	−.13	.02	0	−.05	−.04	.02	0
Anxiety	−.05	−.06	.01	0	.13	.14	.01	5

pational self-direction in explaining the impact of education on values and orientations focuses in both countries on current educational level, since it is current social-stratification position that is at issue. For the United States, occupational self-direction accounts for a substantial portion of education's effects on parental values and on social orientations, with a range of 27% (for parental valuation of self-direction) to 42% (for trustfulness). For Poland, occupational self-direction accounts for even more of current education's effects on values and on social orientations. Thus, in both countries, the role of occupational self-direction in explaining the effects of education both on parental values and on social orientations is greater than Kohn and Schooler believed it to be.[7]

Self-conceptions again provide the one sharp contrast between the United States and Poland. In the U.S., occupational self-direction largely explains the impact of education on idea-conformity, self-confidence, and self-deprecation, albeit not on anxiety; but education in any case has little effect on anxiety. In Poland, occupational self-direction explains little or none of the effect of education on any aspect of self-conception except idea-conformity.

[7] Kohn and Schooler attributed the effects of education on values, self-conceptions, and social orientations mainly to the increased intellectual flexibility and breadth of perspective afforded by educational experience rather than to occupational self-direction (see Kohn and Schooler, 1969:675–7; Kohn, 1969:Chap. 11; see also the extended discussions in Kohn, 1977 and 1981). We defer consideration of this part of their interpretation until we have completed a comparative analysis of job conditions and intellectual flexibility.

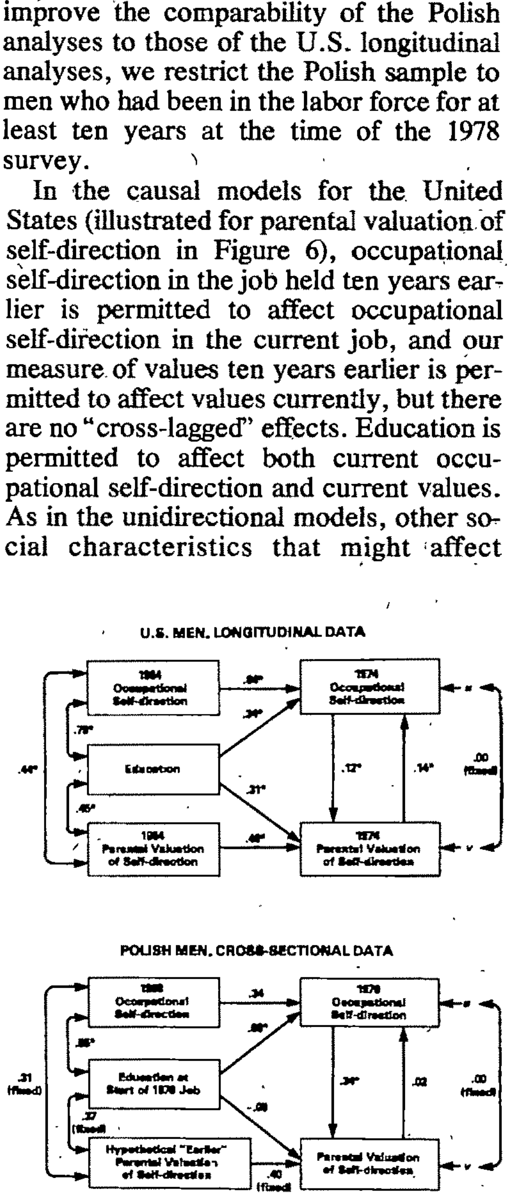
In short, in Poland even more than in the United States, the effects of social-stratification position on parental valuation of self-direction and on social orientations result substantially from men of higher social-stratification position exercising greater self-direction in their work. But, although occupational self-direction similarly accounts for much of the relationship of social-stratification position with idea-conformity, self-confidence, and self-deprecation in the U.S., it fails to do so in Poland.

THE RECIPROCAL EFFECTS OF OCCUPATIONAL SELF-DIRECTION AND PSYCHOLOGICAL FUNCTIONING

Thus far, we have implicitly assumed that values and orientations are affected by, but do not themselves affect, occupational self-direction. Now, following Kohn and Schooler's reformulation of the problem (1973, 1978, 1981, forthcoming; J. Miller et al., 1979; J. Miller, 1980; K. Miller and Kohn, 1980), we abandon the assumption of unidirectionality, recognizing that, as a result of both job-selection and job-molding processes, men's values, self-conceptions, and social orientations may affect their conditions of work. Since a thorough appraisal requires longitudinal data (see Kohn and Schooler, 1978), our ability to assess the reciprocal effects of occupational self-direction and psychological functioning is somewhat constrained by the lack of longitudinal data for Poland. We can, however, make a longitudinal assessment for the United States and, borrowing from the methods developed in J. Miller et al. (1979), we can develop simulated longitudinal models that provide tolerably good approximations for Poland.

For the United States, we use the data of the 1974 follow-up study, based on a representative subsamble of 687 of the men in the original sample. The indices of occupational self-direction, parental values, self-conceptions, and social orientations are based on measurement models that are over-time analogs of the cross-sectional models presented earlier in this paper. (The measurement model for occupational self-direction is developed in Kohn and Schooler, 1981; the model for parental values, in Kohn and Schoenbach, 1980; and those for self-conceptions and social orientations, in Kohn and Schooler, forthcoming.) For Poland, where our only body of data is that of the 1978 survey, we use the indices developed in this paper. To improve the comparability of the Polish analyses to those of the U.S. longitudinal analyses, we restrict the Polish sample to men who had been in the labor force for at least ten years at the time of the 1978 survey.

In the causal models for the United States (illustrated for parental valuation of self-direction in Figure 6), occupational self-direction in the job held ten years earlier is permitted to affect occupational self-direction in the current job, and our measure of values ten years earlier is permitted to affect values currently, but there are no "cross-lagged" effects. Education is permitted to affect both current occupational self-direction and current values. As in the unidirectional models, other social characteristics that might affect



*Statistically significant, $p \leq .05$.

Figure 6. Reciprocal Effects of Occupational Self-direction and Parental Valuation of Self-direction for U.S. and Polish Men (Standardized coefficients. Paths from background characteristics to endogenous variables not shown.)

values are permitted to have such effects; a subset of these social characteristics, those that might be seen by an employer as credentials for a job, are also allowed to affect occupational self-direction.

The causal models for Poland (illustrated in Figure 6) are similar to those for the United States, with two major exceptions. First, the Polish data about the occupational self-direction of the ten-year-earlier job are again limited and retrospective; we use retrospective accounts of substantive complexity as a proxy for the occupational self-direction of the prior job. Second, and much more important, we have no data about ten-year-earlier parental values, self-conceptions, and social orientations. We attempt to deal with this deficiency in two ways. As a baseline, we assess models that simply do not include earlier values and orientations. We then create models in which we simulate a path from a hypothetical earlier to an actual current facet of values or orientation.

The U.S. models (see Table 7) show that occupational self-direction affects parental values and most aspects of orientation. Specifically, occupational self-direction significantly affects parental valuation of self-direction, standards of morality, trustfulness, idea-conformity, self-deprecation, and anxiety. The models also show that values and most aspects of orientation affect occupational self-direction. As hypothesized, the relationships are truly reciprocal.

Since these models do not simultaneously test the reciprocal contemporaneous and the cross-lagged effects of occupational self-direction and values or orientations, we cannot say whether the effects of occupational self-direction on values and orientations are ongoing and continuous or occur only after some lapse of time. Nor can we say whether the effects of values and orientations on occupational self-direction result primarily from men molding their jobs to fit their

Table 7. The Reciprocal Effects of Occupational Self-direction ("Occ.") and Values and Orientations ("Psych."), for Polish and American Men

	Standardized Path Coefficients							
	U.S. Men		Polish Men					
	Earlier Psychological States Controlled		Earlier Psychological States Not Controlled		Hypothetical Stability = .40		Hypothetical Stability = .70	
	Occ. to Psych.	Psych. to Occ.	Occ. to Psych.	Psych. to Occ.	Occ. to Psych.	Psych. to Occ.	Occ. to Psych.	Psych. to Occ.
Values								
Parental valuation of self-direction	.12*	.14*	.57*	−.03	.34*	.02	.18	.06
Social Orientation								
Authoritarian conservatism	−.08	−.18*	−.39*	−.03	−.26*	−.06	−.18	−.08*
Standards of morality	.23*	.24*	.38*	−.03	.23*	.01	.11	.03
Trustfulness	.32*	−.02	.85*	−.06	.52*	.02	.27*	.07*
Self-conception								
Idea conformity	−.22*	−.19*	−.31*	.08*	−.19	.06*	−.09	.04
Self-confidence	.10	.14*	−.19	.05	−.12	.04	−.06	.02
Self-deprecation	−.39*	−.17*	.00	−.01	.00	−.01	.00	−.01
Anxiety	−.27*	−.01	−.19	.03	−.11	.01	−.05	.00

NOTE: The U.S. data are based on a 10-year longitudinal subsample of 687 men, except for parental values, for which the N is 399. The causal models control earlier (1964) psychological states. The correlation of the residuals of the 1974 psychological variable and 1974 occupational self-direction is fixed at zero only if found to be statistically nonsignificant.

The Polish data are based on men interviewed in 1978 who were employed 10 years earlier, 1232 men, except for parental values for which the N is 589. Causal models presented before and after controlling for earlier "hypothetical" psychological states. The correlation of the residuals of the 1978 psychological construct and 1978 occupational self-direction is fixed at zero.

* Statistically significant, $p \leq .05$.

values (a contemporaneous process) or from selection of men into jobs for which they are suited and out of jobs for which they are not suited (a lagged process). Whatever the timing, the relationship of occupational self-direction to values and orientations is quintessentially reciprocal: Of crucial importance for the hypotheses we are testing. Occupational self-direction actually does have substantial causal effects on values and orientations.

For Poland, models that do not include a simulated measure of "earlier" psychological functioning (Table 7) suggest a significant effect of occupational self-direction on parental values and on all three aspects of social orientation. Not surprisingly, occupational self-direction significantly affects only one of the four facets of self-conception, idea-conformity. Thus, in Poland as in the United States, occupational self-direction appears to result in self-directed values and social orientations. However, in Poland the effects of occupational self-direction on self-conceptions are in the main statistically nonsignificant. Finally, there is little evidence in the Polish models that values and orientations affect occupational self-direction.

Since these models do not control "earlier" values and orientations, the findings may exaggerate the effects of occupational self-direction on values and orientations and underestimate the reciprocal effects. We can partially compensate not having information about "earlier" values and orientations by simulating such information. We do this by allowing paths from hypothetical earlier to actual current values and orientations (see Table 7). We fix the hypothetical "stabilities" of each psychological variable at several alternative values, from .40 to .70, the range found in our analyses of the U.S. data.[8] Introduction of these hypothetical stabilities reduces the magnitudes of the paths from occupational self-direction to parental values and to social orientations, to the point where all except the path to trustfulness cease to be statistically significant when the hypothetical stabilities are set at the extreme upper end of the range. Still, all the paths remain substantial, with magnitudes similar to those found in the longitudinal models for the United States, and are consistent with the interpretation that occupational self-direction does affect values and social orientations. At the higher stabilities, there is evidence, too, of social orientations significantly affecting occupational self-direction.

In sum, the longitudinal U.S. models demonstrate a causal impact of occupational self-direction on values and social orientations and the cross-sectional Polish models give consistent results. The one real difference between the U.S. and Polish models is that occupational self-direction significantly affects some aspects of self-conception in the United States but not in Poland.

DISCUSSION

Insofar as cross-national analyses of social structure and personality yield similar findings in the countries studied, our interpretation can ignore whatever differences there may be in the cultures, political and economic systems, and historical circumstances of the particular countries, to deal instead with social-structural universals. But when the relationships between social structure and personality differ from country to country, then we must look to what is idiosyncratic about the particular countries for our interpretation.

[8] To simulate a hypothetical "stability" for some facet of values and orientation, it is necessary to make certain assumptions about the correlations between the hypothetical earlier facet of values and orientation with background characteristics and with occupational self-direction in the earlier job. We base these assumptions on the correlations among current education, age, urbanness of principal place raised, father's status, current occupational self-direction, and current values and orientations. We initially assume that similar associations existed at the earlier time. This assumption is most questionable for experiences that would have been more proximate or less proximate at the earlier time—in particular, the correlations of earlier psychological states with father's status and urbanness of background (more proximate) and with 1978 education (less proximate). We therefore tested the sensitivity of the models to the magnitudes of these correlations by increasing or decreasing them (depending on proximity) by 10%. The conclusions drawn from the reciprocal models are essentially unchanged.

In this inquiry, we find social-structural universals in the pattern of interrelationship among social stratification, occupational self-direction, values, and social orientations. We confirm, for both Poland and the United States, that higher social-stratification position is associated with valuing self-direction and with holding social orientations consonant with valuing self-direction—namely, a nonauthoritarian, open-minded orientation, personally responsible standards of morality, and trustfulness. Our interpretation of these associations need not, indeed should not, focus on anything unique to the United States or to Poland. An adequate interpretation must transcend the two countries, which means, inter alia, that it must apply equally well in both capitalist and socialist society. We seek a general explanation of the psychological ramifications of social stratification in industrialized society. This explanation is readily found in Kohn and Schooler's interpretation that social stratification is associated with values and with social orientations in large part because men of higher position have greater opportunity to be self-directed in their work—to work at substantively complex tasks, free of close supervision, and not subject to routinization.

We not only confirm the findings that underlie this interpretation, both in the United States and in Poland, but also reformulate and test parts of the interpretation for which Kohn and Schooler advanced only an a priori argument. Thus, we reconceptualize and empirically assess the reciprocal relationship between social-stratification position and occupational self-direction. We reassess the extent to which the effects of social stratification on values and orientations are indirect, via occupational self-direction. We also demonstrate that occupational self-direction has an actual causal impact on values and social orientations. By so doing, we provide solid evidence for the interpretation that self-direction in one's work leads to valuing self-direction more generally—as measured here, it leads to valuing self-direction for one's children—and to having a more open, flexible orientation to society. Lack of opportunity for self-direction in one's work leads one to value conformity to external authority and to view social reality as hostile and threatening. In short, the lessons of the job are generalized to one's stance toward the larger society, in socialist as well as in capitalist society. This does not mean that these processes are necessarily the same in all socialist and all capitalist societies, but it does mean that the original U.S. findings are not restricted to capitalist countries.

The one noteworthy difference between Poland and the United States is in the relationship between social-stratification position and self-conception, and in particular, self-confidence. Although the associations between social stratification and self-conceptions in the United States are of only modest magnitude, the pattern is consistent with those for values and social orientations—men of higher social position have more favorable conceptions of self. As with values and social orientations, the relationships between social stratification and self-conceptions result largely from men of higher position being more self-directed in their work. In Poland, there is not the same association between social stratification and self-conceptions; in fact, men of higher social position are somewhat less self-confident and more anxious than are men of lower social position. Nor does occupational self-direction have much impact on self-conceptions. Somehow, in the United States, the learning-generalization process by which the lessons of the job are carried over to one's values and social orientations occurs also (albeit, not as strongly) for one's orientations to self. In Poland, either occupational self-direction does not enhance self-conception or something interferes with such enhanced self-conception being carried over to nonoccupational spheres of life. Moreover, even without occupational self-direction enhancing the self-conceptions of more highly placed men, why does higher social-stratification position fail to have a favorable effect on self-conception, particularly on self-confidence, in Poland? We see several alternative, albeit not incompatible, explanations.

One possibility is that in the aftermath of World War II and the rapid reindustrialization of Poland, many people of working-class and peasant backgrounds who might not otherwise have had the opportunity for higher education and responsible jobs did get these opportunities. This resulted from several historical processes: the Nazis' systematic massacre of the Polish intelligentsia, the rapid increase in industrialization and bureaucratization from pre-war levels, and the deliberate policy of socialist governments to make educational opportunities available to the children of workers and peasants. Whatever the reasons, it is possible that rapid social mobility has resulted in some feelings of self-doubt among people who occupy higher positions. Alternatively, during the transitional stages to a new economic and political system, higher positions may be more precarious than lower positions. As a consequence, people in higher positions may wonder whether they really are sufficiently competent to carry out the responsibilities of their jobs. Another possible explanation focuses on the widespread belief, in part a derivative of the aristocratic culture of pre-war Poland, that status is not so much a result of superior job performance as of fate or the intervention of well-placed friends. Hence, finding oneself in a high position would provide little assurance of one's ability to meet the demands of the position.

These explanations focus on the reasons why Polish men of higher educational and occupational position may not have as favorable self-conceptions as one might have expected. It is equally pertinent to ask why Polish men of lower educational and occupational position may be more self-confident and less anxious than one might have expected. Although, as individuals, these men may not have experienced much social mobility, certainly the segment of the society to which they belong—the working class—has benefitted from post-war changes in socialist Poland. Their economic situation has improved and they are now held in higher social regard; there is every reason for them to feel more confident. Post-war changes in Polish social structure have enhanced the circumstances of the working class to a much greater extent than those of white-collar workers and professionals.

We must emphasize, once again, that occupational self-direction—and, largely as a result, social-stratification position—have much stronger effects on values and on social orientations than on self-conceptions, both in Poland and in the United States. Self-conceptions are much more influenced by nonoccupational, nonstratification facets of life experience. This may be especially so in more traditional societies. Although Poland is an industrial and urbanized nation, traditions emphasizing family and community life, reinforced by religious traditionalism, may be more strongly maintained in Poland than in the United States. In any case, the relationships of stratification and occupational self-direction to self-conceptions clearly are contingent on the cultural, social-structural, and historical circumstances of the countries studied. The relationships of social stratification and job conditions to values and social orientations, though, are stronger and seem to be invariant.

REFERENCES

Alwin, Duane F. and David J. Jackson
1981 "The statistical analysis of Kohn's measures of parental values." In K. G. Joreskog and H. O. A. Wold (eds.), Systems Under Indirect Observation: Causality, Structure, and Prediction. New York: Elsevier North Holland.

Bertram, Hans
1976a "Gesellschaftliche und familiare bedingungen moralischen urteilens." Doctoral dissertation, Universität Duesseldorf.
1976b "Probleme einer sozialstrukturell orientierten Sozialisationsforschung." Zeitschrift für Soziologie 5(April):103–17.

Coburn, David and Virginia L. Edwards
1976 "Job control and child-rearing values." Canadian Review of Sociology and Anthropology 13:337–44.

Danilowicz, Pawel and Pawel Sztabinski
1977 Pytania Metryczkowe. Warsaw: Institute of Philosophy and Sociology of the Polish Academy of Sciences.

Firkowska, Anna, Antonina Ostrowska, Magdalena Sokolowska, Zena Stein, Mervyn Susser, and Ignacy Wald
1978 "Cognitive development and social policy: the contribution of parental occupation and education to mental performance in 11-year olds in Warsaw." Science 200:1357–62.

Fox, John
1980 "Effect analysis in structural equation models." Sociological Methods and Research 9:3–28.

Grabb, Edward G.
1981 "Class, conformity, and political powerlessness." Canadian Review of Sociology and Anthropology (In press).

Heise, David R.
1975 Causal Analysis. New York: John Wiley and Sons.

Hoff, Ernst-Hartmut and Veronika Grueneisen
1977 "Arbeitserfahrungen, erziehungseinstellungen und erziehungsverhalten von eltern." Pp. 65–89 in K. Schneewind und H. Lukesch (eds.), Familiare Sozialisation: Probleme, Ergebnisse, Perspektiven. Stuttgart: Klett-Cotta.

Hynes, Eugene
1977 Personal communication (Unpublished data).

Jackson, David J. and Duane F. Alwin
1980 "The factor analysis of ipsative measures." Sociological Methods and Research 9:218–38.

Joreskog, Karl G.
1969 "A general approach to confirmatory maximum likelihood factor analysis." Psychometrika 34:183–202.

Joreskog, Karl G. and Marielle van Thillo
1972 "LISREL: a general computer program for estimating a linear structural equation system involving multiple indicators of unmeasured variables." Research Bulletin 72–56. Princeton, N.J.: Educational Testing Service.

Kohn, Melvin L.
1963 "Social class and parent-child relationships: an interpretation." American Journal of Sociology 68:471–80.
1969 Class and Conformity: A Study in Values. Homewood, Ill.: Dorsey Press. (Second edition, University of Chicago Press, 1977).
1977 "Reassessment, 1977." Pp. xxv-lx in Class and Conformity: A Study in Values. Second edition. Chicago: University of Chicago Press.
1981 "Personality, occupation, and social stratification: a frame of reference." Pp. 267–97 in Donald J. Treiman and Robert V. Robinson (eds.), Research in Social Stratification and Mobility. Vol. 1. New York: JAI Press.

Kohn, Melvin L. and Carrie Schoenbach
1980 "Social stratification and parental values: a multi-national assessment." Paper presented to the Japan-U.S. Conference on Social Stratification and Mobility. Maui, Hawaii, January.

Kohn, Melvin L. and Carmi Schooler
1969 "Class, occupation, and orientation." American Sociological Review 34:659–78.
1973 "Occupational experience and psychological functioning: an assessment of reciprocal effects." American Sociological Review 38:97–118.
1978 "The reciprocal effects of the substantive complexity of work and intellectual flexibility: a longitudinal assessment." American Journal of Sociology 84:24–52.
1981 "Job conditions and intellectual flexibility: a longitudinal assessment of their reciprocal effects." Pp. 281–313 in David J. Jackson and Edgar F. Borgatta (eds.), Factor Analysis and Measurement in Sociological Research: A Multi-Dimensional Perspective. London: Sage Publications.
forthcoming "Job conditions and personality: a longitudinal assessment of their reciprocal effects." American Journal of Sociology.

Lawley, D. N. and A. E. Maxwell
1971 Factor Analysis as a Statistical Method. 2nd Ed. London: Butterworth.

Miller, Joanne
1980 "Individual and occupational determinants of job satisfaction: a focus on gender differences." Sociology of Work and Occupations 7:337–66.

Miller, Joanne, Carmi Schooler, Melvin L. Kohn, and Karen A. Miller
1979 "Women and work: the psychological effects of occupational conditions." American Journal of Sociology 85:66–94.

Miller, Joanne, Kazimierz Slomczynski, and Ronald Schoenberg
1981 "Assessing comparability of measurement in cross-national research: authoritarian-conservatism in different sociocultural settings." Social Psychology Quarterly, in press.

Miller, Karen A. and Melvin L. Kohn
1980 "The reciprocal effects of job conditions and the intellectuality of leisure-time activity." Paper presented to the Japan-U.S. Conference on Social Stratification and Mobility. Maui, Hawaii, January.

Mortimer, Jeylan T. and Jon Lorence
1979a "Work experience and occupational value socialization: a longitudinal study." American Journal of Sociology 84:1361–85.
1979b "Occupational experience and the self-concept: a longitudinal study." Social Psychology Quarterly 42:307–23.

Nowak, Stefan
1969 "Changes of social structure in social consciousness." Pp. 235–47 in Celcia Heller (ed.), Structured Social Inequality: A Reader in Comparative Social Stratification. New York: Macmillan.

Pohoski, Michal, Kazimierz Slomczynski, and Wlodzimierz Wesolowski
1976 "Occupational prestige in Poland, 1958–1975." Polish Sociological Bulletin, No. 4:86–94.

Rosenberg, Morris
1979 Conceiving the Self. New York: Basic Books.

Sarapata, Adam, and Wlodzimierz Wesoloski
1961 "The evaluation of occupations by Warsaw inhabitants." American Journal of Sociology 66:581–91.

Schooler, Carmi
1980 "Psychological and social perspectives on

status attainment." Paper presented to the Japan-U.S. Conference on Social Stratification and Mobility. Maui, Hawaii, January.

Slomczynski, Kazimierz M., and Grazyna Kacprowicz
1979 Skale Zawodow. Warsaw: Institute of Philosophy and Sociology of the Polish Academy of Sciences.

Slomczynski, Kazimierz M. and Tadeusz Krauze (eds.)
1978 Class Structure and Social Mobility in Poland. White Plains, N.Y.: M. E. Sharpe.

Telenbach, Sten
1974 "Patterns of stratification in socialist Poland." Acta Sociologica 17:25–47.

Treiman, Donald J.
1977 Occupational Prestige in Comparative Perspective. New York: Academic Press.

Wesolowski, Wlodzimierz
1969 "The notions of strata and class in socialist society." Pp. 122–45 in Andre Beteille (ed.), Social Inequality: Selected Readings. Harmondsworth, Middlesex, England: Penguin Books.
1975 Teoria, Badania, Praktyka. Warszawa: Ksiazka i Wiedza.
1979 Classes, Strata and Power. London: Routledge and Kegan Paul.

Wesolowski, Wlodzimierz, and Kazimierz M. Slomczynski
1977 Investigations on Class Structure and Social Stratification in Poland, 1945–1975. Warsaw: The Institute of Philosophy and Sociology of the Polish Academy of Sciences.

BLACK SOUTHERN STUDENT SIT-IN MOVEMENT: AN ANALYSIS OF INTERNAL ORGANIZATION*

ALDON MORRIS
University of Michigan

This paper argues that the Southern sit-in movement of 1960, though it appears to have developed in the spontaneous manner described by classic collective behavior theory, actually grew out of pre-existing institutions and organizational forms. The spread of the sit-ins followed the networks of these pre-existing institutional relationships. Factors internal to the black community—churches, colleges, protest organizations, and leaders—were responsible for nurturing and developing the movement. The analysis is based on primary data collected from archives and interviews with civil rights leaders.

Scholars of the Civil Rights movement (Zinn, 1964; Oppenheimer, 1964; Matthews and Prothro, 1966; Meier and Rudwick, 1973; Oberschall, 1973; McAdam, 1979) and Civil Rights activists agree that the black Southern student sit-in movement of 1960 was a crucial development. The sit-ins pumped new life into the Civil Rights movement and enabled it to win unprecedented victories. Moreover, the sit-ins exercised a profound tactical and strategic influence over the entire course of social and political upheavals of the 1960s.

Apart from having a jarring impact on race relations, the sit-ins signaled the possibility of militant action at both Northern and Southern white campuses (Haber, 1966; Obear, 1970; Sale, 1973). A critical mass of the early leaders of the white student movement acquired much of their training, organizing skills, and tactics from the black activists of the student sit-in movement (Sale, 1973; Westby,

* Direct all communications to Aldon Morris, Center for Research on Social Organization, University of Michigan, 330 Packard, Ann Arbor, Michigan 48109.

I would like to thank Kim Myles, Walter Allen, Michael Schwartz, Charles Perrow, Lewis Coser, Doug McAdam, Mayer Zald, William Gamson, and Charles Tilly for their helpful comments on this paper. The debt that I owe movement participants who consented to be interviewed will be obvious. A special thanks to Sheila Wilder and Debbie Snovak who labored through several drafts of this paper. Finally I thank two anonymous ASR Reviewers for extremely valuable comments. This research was partially supported by the ASA Minority Fellowship Program and a grant from the National Science Foundation SOC 76-20171.

1976). Thus, the beginning of the white student movement as well as the quickened pace of Civil Rights activity can be traced to the black student sit-in movement.

The sit-ins were important because their rapid spread across the South crystalized the conflict of the period and pulled many people directly into the movement. How is such a "burst" of collective action to be explained? A standard account of the sit-ins has emerged which maintains that the sit-ins were the product of an independent black student movement which represented a radical break from previous civil rights activities, organizations, and leadership of the Black community (e.g. Lomax, 1962; Zinn, 1964; Oppenheimer, 1964; Matthews and Prothro, 1966; Meier and Rudwick, 1973; Oberschall, 1973; Piven and Cloward, 1977).

In the standard account, various factors are argued to be the driving force behind the sit-ins, including impatience of the young, mass media coverage, outside resources made available by the liberal white community of the North, and support from the Federal Government. Although these writers differ over the proximate causes of the sit-ins, they nevertheless concur that the sit-ins broke from the organizational and institutional framework of the emerging Civil Rights movement. The data for the present study do not fit this standard account and suggest that a different account and interpretation of the sit-ins is warranted. The purpose of this paper is to present new data on the Southern student sit-in movement of 1960, and to provide a framework that will theoretically order the empirical findings.

THEORETICAL CONTEXT AND PROPOSITIONS

Classical collective behavior theory and the recently formulated resource mobilization theory are the major sociological frameworks that attempt to provide explanations of the origins, development, and outcomes of social movements. Classical collective behavior theory (e.g. Blumer, 1946; Turner and Killian, 1957; Lang and Lang, 1961; and Smesler, 1963) maintains that social movements differ substantially from institutionalized behavior. Social movements are theorized to be relatively spontaneous and unstructured. Movement participants are often portrayed as nonrational actors functioning outside of normative constraints and propelled by high levels of strain.

Classical collective behavior theorists do not deny that organizations and institutional processes play a role in collective behavior. Rather, organizations and institutional processes emerge in the course of movements and become important in their later stages. The standard account of the sit-ins fits the collective behavior imagery. Indeed, it can be argued that the diverse proponents of the "standard account" have been unduly influenced by classical collective behavior theory; their account largely ignores the organizational and institutional framework out of which the sit-ins emerged and spread.

The resource mobilization explanation (e.g. Oberschall, 1973; Gamson, 1975; Tilly, 1978; McCarthy and Zald, 1973) of social movements differs markedly from classical collective behavior theory. In this view, social movements have no distinct inner logic and are not fundamentally different from institutionalized behavior. Organizations, institutions, pre-existing communication networks, and rational actors are all seen as important resources playing crucial roles in the emergence and outcome of collective action. In contrast to classical collective behavior theory, organizational and institutional structures are argued to be central throughout the entire process of collective action.

In its present formulation, resource mobilization theory is unclear about the type of organization and resources that are crucial for the initiation and spread of collective action. Some theorists (Oberschall, 1973; McCarthy and Zald, 1973; Jenkins and Perrow, 1977) argue that resources and organizations outside the protest group are crucial in determining the scope and outcomes of collective action. External groups and resources are argued to be especially critical for movements of the poor. In other formulations

of this approach (e.g. Gamson, 1975; Tilly, 1978), emphasis is placed on the important role that internal organization plays in collective action. However, internal organization is but one of several variables (e.g. repression, bureaucracy, opportunity) that are investigated. In my view such an approach fails to capture the degree to which collective action is dependent on internal organization.

This paper focuses on the central function that internal organization played in the emergence and development of the sit-in movement. My analysis suggests that one-sided emphases on spontaneous processes or outside resources can lead to unwarranted neglect of internal structure. A case will be made that the diffusion of the 1960 sit-ins cannot be understood without treating internal organization as a central variable. The analysis will be guided by three propositions.

Proposition 1. Pre-existing social structures provide the resources and organizations that are crucial to the initiation and spread of collective action. Following Tilly (1978), collective action is defined here as joint action by protest groups in pursuit of common ends. This proposition maintains that collective action is rooted in organizational structure and carried out by rational actors attempting to realize their ends. This proposition is central to resource-mobilization theory and has received considerable support from a number of empirical studies (Oberschall, 1973; Gamson, 1975; Tilly, 1975).

Proposition 2. The extent and distribution of internal social organization will determine the extent to which innovations in collective strategy and tactics are adopted, spread, and sustained. This proposition directs attention to a protest group's internal organization—its "local movement centers." A local movement center is that component of social structure within a local community that organizes and coordinates collective action. A local movement center has two major properties. First, it includes all protest organizations and leaders of a specific community that are actively engaged in organizing and producing collective action. During the sit-ins, the Southern Christian Leadership Conference (SCLC), Youth Councils of the National Association for the Advancement of Colored People (NAACP), Congress of Racial Equality (CORE), and "direct action" churches existed in numerous Southern black communities. A local center within the Civil Rights movement included all these organizations and leaders. Second, a local movement center contains a unit that coordinates protest activities within the local movement and between the local center and other institutions of the larger community. During the Civil Rights movement, a particular church usually served as the local coordinating unit. Through this unit the protest activities of the church community, college community, activist organizations, and their leaders were mobilized and coordinated. Thus, movement centers provide the organization and coordination capable of sustaining and spreading collective action.

Proposition 3. There is an interaction between the type of pre-existing internal organization and the type of innovations in strategy and tactics that can be rapidly adopted and spread by a protest group. This proposition addresses the issue of why a protest group adopts a particular tactical innovation rather than another.[1] Whereas Proposition II maintains that diffusion of an innovation in strategy is a function of the development and spread of internal social organization, Proposition III specifies that certain types of organization are more conducive than others to the diffusion and adoption of certain types of tactical innovation.

In short, the framework for the analysis of the 1960 sit-ins consists of three interrelated propositions. One, collective action is initiated through pre-existing

[1] Why, for example, did the "teach-ins" spread rapidly between college campuses during the mid-sixties? This proposition suggests that the teach-in tactic was especially suited to the university-based internal organization of the white student movement. In its essentials the teach-in innovation was academically oriented and could be implemented by academic types who were entrenched in the "movement centers" of the various universities involved in the movement. Lecture halls, libraries, film clips, study groups, seminar notes, etc. were the pre-existing indigenous resources used by agents of the movement via the teach-ins.

structures. Two, tactical innovation within a movement is a function of well-developed and widespread internal organization. Three, the type of innovation in strategy and tactics which can be rapidly disseminated and sustained is largely determined by the characteristic internal organization of a protest group.

DATA

This study of the sit-ins is part of a larger study on the origins of the Civil Rights movement (Morris, forthcoming). A substantial part of the data were collected from primary sources—archives and interviews with Civil Rights participants. The archival research was conducted at various sites between May and September of 1978.[2] Thousands of original documents (i.e. memoranda, letters, field reports, organizational histories and directives, interorganizational correspondences, etc.) generated by movement participants were examined. These data contained a wealth of information pertaining to key variables—organization, mobilization, finance, rationality, spontaneity—relevant to the study of movements.

Interviews with participants of the movement constituted the second source of data. Detailed interviews with over 50 Civil Rights leaders were conducted. Interviews made it possible to follow-up on many issues raised by the archival data; and, since these interviews were semi-open-ended, they revealed unexpected insights into the movement. Whenever statements were heard that seemed novel or promising, interviewees were given freedom to speak their piece.

METHODS

The strategy for the archival research was straightforward. The researcher examined every document possible within the time allocated for a particular site.[3] The main objective was to examine the roles played in the sit-ins by variables associated with Weberian theory and theories of collective behavior and resource mobilization. Following collective behavior theory, I was concerned with the extent to which the sit-ins were spontaneous and discontinuous with established social structure. From Weberian theory I was interested in whether a charismatic attraction between a leader and followers was sufficient to produce the heavy volume of collective action in the 1960 sit-ins. Finally, several issues addressed by resource mobilization theory were of interest. I examined archival sources to ascertain the role of social organization and resources in the sit-ins. Also, I was concerned with whether the leadership, money, and skills behind the sit-ins were supplied by outsiders or by the indigenous Southern black community.

Three strategies were employed in the interview process. First, the researcher attempted to learn as much as possible about the movement from extensive library and archival sources before conducting interviews. This prior knowledge enabled the interviewer to ask specific questions and to assist interviewees in rooting their memories in the social, temporal, and geographical context of their actions twenty years earlier. Prior knowledge enabled the interviewer to gain the respect of interviewees and increased the likelihood that they would approach the interview in a serious manner.

Second, the interviews were semi-structured, usually lasting two or three hours. An extended list of questions structured around the variables used in the archival research were formulated beforehand. The interviewees were instructed to feel free to deviate from the questions and to discuss what they thought to be important. Their "diversions" produced new information.

Third, the interview sample was assembled in two ways. While examining the archival material, the names of leaders associated with various activities turned up constantly. These were the initial individuals contacted for interviews. Once the interview process was underway, interviewees would invariably remark, often in

---

[2] King papers at Boston University; SCLC papers at the Southern Christian Leadership Conference headquartered in Atlanta; Rev. Kelly Miller Smith's papers housed at First Baptist Church of Nashville.

[3] All of the King papers at Boston University and all of SCLC's files in Atlanta were examined, as well as the portion of Rev. Smith's papers dealing with the sit-ins.

response to queries, "you know, you really should speak to [so-and-so] regarding that matter." Subsequent interviews were arranged with many of these individuals. Thus, the snowball effect was central to the sampling process. Although the activists interviewed came from numerous organizations and represented different, if not conflicting, viewpoints, to our surprise they agreed on many basic issues.

Given that the sit-in movement occurred twenty years ago, it is reasonable to wonder whether interview accounts are reliable and valid. Moreover, there is the suspicion that participants might have vested interests in presenting the "facts" in such a way as to enhance their own status. Such problems of recall and vested interest have been minimized in this research because the analysis is not based on any one source. Rather, it is built on an array of published material, archival sources, and accounts of individuals who participated in and were eye-witnesses to the same events. Furthermore, cross references were made throughout the data collection process. Follow-up phone calls were made to clarify ambiguity and to obtain a comprehensive view of the sit-in movement. It appears that neither of these potential trouble spots produced fundamental defects in the data.

EARLY SIT-INS: FORERUNNERS

The first myth regarding the sit-in movement is that it started in Greensboro, North Carolina, on February 1, 1960. This research documents that Civil Rights activists conducted sit-ins between 1957 and 1960 in at least fifteen cities: St. Louis, Missouri; Wichita and Kansas City, Kansas; Oklahoma City, Enid, Tulsa, and Stillwater, Oklahoma; Lexington and Louisville, Kentucky; Miami, Florida; Charleston, West Virginia; Sumter, South Carolina; East St. Louis, Illinois; Nashville, Tennessee; and Durham, North Carolina.[4] The Greensboro sit-ins are important because they represent a unique link in a long chain of sit-ins. Although this paper concentrates on the uniqueness of the Greensboro link, there were important similarities in the entire chain. While other studies (Southern Regional Council, 1960; Oppenheimer, 1964; Matthews and Prothro, 1966; Meier and Rudwick, 1973) have not totally overlooked these earlier sit-ins, they fail to reveal their scope, connections, and extensive organizational base.

The early sit-ins were initiated by direct-action organizations. From interviews with participants in the early sit-ins (Moore, 1978; McCain, 1978; Lawson, 1978; Smith, 1978; McKissick, 1978, 1979; Luper, 1981; Randolph, 1981; Lewis, 1981) and published works (Southern Regional Council, 1960; Meier and Rudwick, 1973), I found that Civil Rights organizations initiated sit-ins in fourteen of the fifteen cities I have identified. The NAACP, primarily its Youth Councils, either initiated or co-initiated sit-ins in nine of the fifteen cities. CORE, usually working with the NAACP, played an important initiating role in seven of the fifteen cities. The SCLC initiated one case and was involved in another. Finally, the Durham Committee on Negro Affairs, working with the NAACP, initiated sit-ins in that city. From this data, we can conclude that these early sit-ins were a result of a multi-faceted organizational effort.

These sit-ins received substantial backing from their respective communities. The black church served as the major institutional force behind the sit-ins. Over two decades ago, E. Franklin Frazier argued that "for the Negro masses, in their social and moral isolation in American society, the Negro church community has been a nation within a nation" (Frazier, 1963:49). He argued that the church functioned as the central political arena in black society. Nearly all of the direct-action organizations that initiated these early sit-ins were closely associated with the church. The church supplied these organizations not only with an established communication network, but also leaders and organized masses, finances, and a safe environment in which to hold political meetings. Direct-action organizations clung to the church because their survival depended on it.

[4] I suspect that further research will reveal that sit-ins occurred in more than these fifteen cities between 1957 and 1960.

Not all black churches supported the sit-ins. The many that did often supported sit-ins in a critical but "invisible" manner. Thus, Mrs. Clara Luper, the organizer of the 1958 Oklahoma City sit-ins, wrote that the black church did not want to get involved, but church leaders told organizers, "we could meet in their churches. They would take up a collection for us and make announcements concerning our worthwhile activities" (Luper, 1979:3). This "covert" role was central. Interviewed activists revealed that clusters of churches were usually directly involved with the sit-ins. In addition to community support generated through the churches, these activists also received support from parents whose children were participating in demonstrations.

These sit-ins were organized by established leaders of the black community. The leaders did not spontaneously emerge in response to a crisis, but were organizational actors in the full sense of the word. Some sit-in leaders were also church leaders, taught school, and headed up the local direct-action organization; their extensive organizational linkages provided blocks of individuals to serve as demonstrators. Clara Luper wrote, "The fact that I was teaching American History at Dungee High School in Spencer, Oklahoma and was a member of the First Street Baptist Church furnished me with an ample number of young people who would become the nucleus of the Youth Council" (Luper, 1979:1). Mrs. Luper's case is not isolated; leaders of the early sit-ins were enmeshed in organizational networks and were integral members of the black community.

Rational planning was evident in this early wave of sit-ins. During the late fifties, the Revs. James Lawson and Kelly Miller Smith, both leaders of a direct-action organization—Nashville Christian Leadership Council—formed what they called a "nonviolent workshop." In these workshops, Lawson meticulously taught local college students the philosophy and tactics of nonviolent protest (D. Bevel, 1978; Lewis, 1978).[5] In 1959, these students held "test" sit-ins in two department stores. Earlier, in 1957, members of the Oklahoma City NAACP Youth Council created what they called their "project," whose aim was to eliminate segregation in public accommodations (Luper, 1979:3). The project consisted of various committees and groups who planned sit-in strategies. After a year of planning, this group walked into the local Katz Drug Store and initiated their sit-in. In St. Louis in 1955, William Clay organized an NAACP Youth Council. Through careful planning and twelve months of demonstrations, members of this organization were able to desegregate dining facilities at department stores (Meier and Rudwick, 1973:93). In Durham, North Carolina in 1958, black activists of the Durham Committee on Negro Affairs conducted a survey of 5-and-10-cent stores in Durham (Southern Regional Council, 1960). The survey revealed that these stores were heavily dependent on black trade. Clearly, the sit-ins initiated by this group were based on rational planning. A similar picture emerges in Sumter, South Carolina and for all the early sit-ins.

Finally, these early sit-ins were sponsored by indigenous resources of the black community; the leadership was black, the bulk of the demonstrators were black, the strategies and tactics were formulated by blacks, and the finances came out of the pockets of blacks, while their serene spirituals echoed through the churches.[6]

Most of the organizers of the early sit-ins knew each other and were well aware

[5] Actual names of movement participants are used in this study rather than pseudonyms. I decided to use actual names because my study focuses on real places, movements, and activists. This approach will assist other researchers in evaluating the interview data, since they will know who said what and can conduct further interviews if the need arises. In addition, the respondents had a story to convey and expressed no desires to remain anonymous.

[6] It could legitimately be argued that outside resources were central to these early sit-ins, given that in some cases CORE was involved. However, it seems that the emerging black, direct-action organizations of the late 1950s and the church served as a resource base for CORE. Thus, CORE, which was very small at the time, "piggybacked" on indigenous resources of the black community. Elsewhere (1980) I have presented supporting data for this argument. Meier and Rudwick's account of early CORE suggests a similar conclusion.

of each other's strategies of confrontation. Many of these activists were part of the militant wing of the NAACP. Following the Montgomery bus boycott, this group began to reorganize NAACP Youth Councils with the explicit purpose of initiating direct-action projects. This group of activists (e.g., Floyd McKissick, Daisy Bates, Ronald Walters, Hosea Williams, Barbara Posey, Clara Luper, etc.) viewed themselves as a distinct group, because the national NAACP usually did not approve of their direct-action approach or took a very ambivalent stance.

These militants of the NAACP built networks that detoured the conservative channels and organizational positions of their superiors. At NAACP meetings and conferences, they selected situations where they could present freely their plans and desires to engage in confrontational politics. At these gatherings, information regarding strategies was exchanged. Once acquainted, the activists remained in touch by phone and mail.

Thus, it is no accident that the early sit-ins occurred between 1957 and 1960. Other instances of 'direct action' also occurred during this period. For example, Mrs. Daisy Bates led black students affiliated with her NAACP Youth Council into the all-white Little Rock Central High School and forced President Eisenhower to send in National Guards. CORE, beginning to gain a foothold in the South, had the explicit goal of initiating direct-action projects. We have already noted that CORE activists were in close contact with other activists of the period. Though these early sit-ins and related activities were not part of a grandiose scheme, their joint occurrences, timing, and approaches were connected via organizational and personal networks.

SIT-IN CLUSTER

Organizational and personal networks produced the first cluster of sit-ins in Oklahoma in 1958. By tracing these networks, we can arrive at a basic understanding of this cluster and a clue to understanding the entire sit-in movement.

In August of 1958, the NAACP Youth Council of Wichita, Kansas, headed by Ronald Walters, initiated sit-ins at the lunch counters of a local drug store (Lewis, 1981). At the same time, Clara Luper and the young people in her NAACP Youth Council were training to conduct sit-ins in Oklahoma City. The adult leaders of these two groups knew each other: in addition to working for the same organization, several members of the two groups were personal friends. Following the initial sit-ins in Wichita, members of the two groups made numerous phone calls, exchanged information, and discussed mutual support. This direct contact was important because the local press refused to cover the sit-ins. In less than a week, Clara Luper's group in Oklahoma City initiated their planned sit-ins.

Shortly thereafter, sit-ins were conducted in Tulsa, Enid, and Stillwater, Oklahoma. Working through CORE and the local NAACP Youth Council, Clara Luper's personal friend, Mrs. Shirley Scaggins, organized the sit-ins in Tulsa (Luper, 1981). Mrs. Scaggins had recently lived in Oklahoma City and knew the details of Mrs. Luper's sit-in project. The two leaders worked in concert. At the same time, the NAACP Youth Council in Enid began to conduct sit-ins. A Mr. Mitchell who led that group (Luper, 1981) knew Mrs. Luper well. He had visited the Oklahoma Youth Council at the outset of their sit-in and discussed with them sit-in tactics and mutual support. The Stillwater sit-ins appear to have been conducted independently by black college students.

A process similar to that in Oklahoma occurred in East St. Louis, Illinois. Homer Randolph, who in late 1958 organized the East St. Louis sit-ins, had previously lived in Oklahoma City, knew Mrs. Luper well, and had young relatives who participated in the Oklahoma City sit-ins.

In short, the first sit-in cluster occurred in Oklahoma in 1958 and spread to cities within a hundred-mile radius via established organizational and personal networks. The majority of these early sit-ins were (1) connected rather than isolated, (2) initiated through organizations and personal ties, (3) rationally planned and led by established leaders, and (4) supported by indigenous resources. Thus, the Greensboro sit-ins did not mark the

movement's beginning, but were links in the chain. But the Greensboro sit-ins were a unique link which triggered sit-ins across the South at an incredible pace. What happened in the black community between the late 1950s and early 1960s to produce such a movement?

EMERGENCE OF INTERNAL ORGANIZATION

During the mid-fifties the extensive internal organization of the Civil Rights movement began to crystalize in communities across the South. During this period "direct action" organizations were being built by local activists. Community institutions—especially the black church—were becoming political. The "mass meeting" with political oratory and protest music became institutionalized. During the same period, CORE entered the South with intentions of initiating protest, and NAACP Youth Councils were reorganized by young militant adults who desired to engage in confrontational politics.

However, neither CORE nor the NAACP Youth Councils were capable of mobilizing wide-scale protest such as the sit-ins of 1960, because neither had a mass base in the black community. CORE was small, Northern-based, and white-led, largely unknown to Southern blacks. Historically, the NAACP had been unable to persuade more than 2% of the black population to become members. Furthermore, the national NAACP was oriented to legal strategies, not sit-ins. Following the 1954 school desegregation decision, the NAACP was further weakened by a severe attack by local white power structures. Members of the Southern white power structures attempted to drive local branches of NAACP out of existence by labeling them subversive and demanding they make their membership public. NAACP officials usually refused to comply with this demand because their members might suffer physical and economic reprisals if identified. NAACP's opponents argued in the local courts that this noncompliance confirmed their suspicion that NAACP was subversive, and the courts responded by issuing injunctions which prevented NAACP from operating in a number of Southern states. For example the NAACP was outlawed in the state of Alabama from 1956 to 1965 (Morris, 1980). This repression forced the NAACP to become defensively-oriented and to commit its resources to court battles designed to save itself. Thus, neither CORE nor NAACP Youth Councils were able to provide the political base required to launch the massive sit-ins of 1960.

Nevertheless, between 1955 and 1960 new organizational and protest efforts were stirring in Southern black communities. The efforts attracted CORE southward and inspired the direct-action groups in the NAACP to reorganize its Youth Councils. The Montgomery bus boycott was the watershed. The importance of that boycott was that it revealed to the black community that mass protests could be successfully organized and initiated through indigenous resources and institutions.

The Montgomery bus boycott gave rise to both the Montgomery Improvement Association (MIA) and the Southern Christian Leadership Conference (SCLC). The MIA was organized in December 1955 to coordinate the activities of the mass bus boycott against segregated buses and to serve as the boycott's official decision-making body. The MIA was a local church-based Southern organization. Its leadership was dominated by local ministers of Montgomery, with the Rev. Martin Luther King serving as its first president. The dramatic Montgomery boycott triggered similar boycotts in a number of Southern cities. As in Montgomery, these boycotts were organized through the churches, with a local minister typically becoming the official leader. SCLC was organized in 1957 by activist clergymen from across the South to coordinate and consolidate the various local movements. SCLC's leadership was dominated by black ministers with King elected as its first president, and the major organizational posts were filled by ministers who led local movements. Thus, SCLC was organized to accomplish across the South what the MIA had in Montgomery. The emergence of MIA and SCLC reflected

the dominant role that churches began to play in confrontational politics by the late 1950s.

The Montgomery bus boycott demonstrated the political potential of the black church and church-related direct-action organizations. By 1955 the massive migration of blacks from rural to urban areas was well underway, and many Southern cities had substantial black populations. The black urban churches that emerged in these cities were quite different from their rural counterparts. The urban churches were larger, more numerous, and better financed, and were presided over by ministers who were better educated and whose sole occupation was the ministry (Mays and Nicholson, 1933; McAdam, 1979; Morris, 1980). Moreover, urban churches were owned, operated, and controlled by the black community.

These churches functioned as the institutional base of the Montgomery bus boycott. They supplied the movement with money, organized masses, leaders, highly developed communications, and relatively safe environments where mass meetings could be held to plan confrontations. This institutional base was in place prior to the boycott. Movement leaders transformed the churches into political resources and committed them to the ends of the movement. The new duty of the church finance committee was to collect money for the movement. The minister's new role was to use the pulpit to articulate the political responsibilities of the church community. The new role of the choir was to weave political messages into the serene spirituals. Regular church meetings were transformed into the "mass meeting" where blacks joined committees to guide protests, offered up collections to the movement, and acquired reliable information of the movement, which local radio and television stations refused to broadcast. The resources necessary to initiate a black movement were present in Montgomery and other communities. They were transformed into political resources and used to launch the first highly visible mass protest of the modern Civil Rights movement.

The important role of the MIA in the emergence of the modern Civil Rights movement is seldom grasped. As a nonbureaucratic, church-based organization, MIA's organizational affairs were conducted like church services rather than by rigid bureaucratic rules, as in the case of the NAACP. Ministers presided over the MIA the way they presided over their congregations. Ultimate authority inhered in the president, Dr. King. Decisions pertaining to local matters could be reached immediately. Diverse organizational tasks were delegated to the rank-and-file on the spot. Rules and procedures emerged by trial and error and could be altered when they inhibited direct action. Oratory, music, and charismatic personalities energized MIA's organizational affairs. The structure of the organization was designed to allow masses to participate directly in protest activities. The MIA proved to be appropriate for confrontational politics because it was mass-based, nonbureaucratic, Southern-led, and able to transform pre-existing church resources into political power.

Southern blacks took notice of the Montgomery movement. Activists from across the South visited Montgomery to observe the political roles of the church and the MIA. For example, when Hosea Williams (at the time, an activist associated with the NAACP in Savannah, Georgia) visited the Montgomery movement, he marvelled at its dynamics:

> You had had NAACP lawsuits, you'd had NAACP chapters, who had much less than 5% participation anyplace. But here's a place [Montgomery] where they got masses of blacks—they couldn't get a church big enough where they could hold mass rallies. And then, none of them [masses] were riding the buses. I was interested in these strategies and their implementation and in learning how to mobilize the masses to move in concert. [Williams, 1978]

Williams, like countless others, did more than marvel. In his words, "I went back to Savannah and organized the Youth Council and nonviolent movement." Thus, another direct-action organization emerged.

Black ministers were in the best position to organize church-related direct-action organizations in the South. Even while the Montgomery movement was in progress, ministers in other cities (e.g.,

Steele in Tallahassee, Shuttlesworth in Birmingham, and Davis in New Orleans) began to build mass-based movements patterned after the Montgomery movement. These ministers were not only in a position to organize and commit church resources to protest efforts, they were also linked to each other and the larger community via ministerial alliances. In short, between 1955 and 1960 a profound change in Southern black communities had begun. Confrontational politics were thrust to the foreground through new direct-action organizations closely allied with the church.

SCLC AND MOVEMENT CENTERS

The creation of the Southern Christian Leadership Conference (SCLC) in 1957 marked a critical organizational shift for the Civil Rights movement. The ministers who organized SCLC clearly understood the historic and central institutional importance of the church in black society. They knew that the church nurtured and produced most of the indigenous leaders, raised finances, and organized masses, as well as being a major force in other aspects of black culture. By 1957 these ministers, many of whom were leading movements in their local communities, consciously and explicitly concluded that the church was capable of functioning as the institutional vanguard of a mass-based black movement. Hence, they organized SCLC to be a Southern-wide, church-based protest organization.

Prior to SCLC, the major black protest organization—NAACP—had been closely linked with the church. Yet, before SCLC was created, the NAACP, and not the church, functioned as the organization through which protest was initiated. With the emergence of SCLC, the critical shift occurred whereby the church itself, rather than groups closely linked to it, began to function as the institutional center of protest.

In 1957 the organizers of SCLC sent out a call to fellow clergymen of the South to organize their congregations and communities for collective protest. The remarks of Rev. Smith of Nashville typified the action of protest-oriented ministers:

> After the meeting [SCLC organizing meeting] and after the discussion that we had and all that, it became clear to me that we needed something in addition to NAACP. So I came back and I called some people together and formed what we named the Nashville Christian Leadership Council in order to address the same kind of issues that SCLC would be addressing. [Smith, 1978]

Hundreds of ministers across the South took similar action.

From this collective effort resulted what can best be conceptualized as local movement centers of the Civil Rights movement, which usually had the following seven characteristics:

1. A cadre of social-change-oriented ministers and their congregations. Often one minister would become the local leader of a given center and his church would serve as the coordinating unit.
2. Direct action organizations of varied complexity. In many cities local churches served as quasi-direct-action organizations, while in others ministers built complex, church-related organizations (e.g., United Defense League of Baton Rouge, Montgomery Improvement Association, Alabama Christian Movement for Human Rights of Birmingham, Petersburg Improvement Association). NAACP Youth Councils and CORE affiliates also were components of the local centers.
3. Indigenous financing coordinated through the church.
4. Weekly mass meetings, which served as forums and where local residents were informed of relevant information and strategies regarding the movement. These meetings also built solidarity among the participants.
5. Dissemination of nonviolent tactics and strategies. The leaders articulated to the black community the message that social change would occur only through nonviolent direct action carried out by masses.
6. Adaptation of a rich church culture to political purposes. The black spirituals, sermons, and prayers were used to deepen the participants' commitment to the struggle.
7. A mass-based orientation, rooted in the black community, through the church.

See Figure 1 for a schematic diagram of a typical local movement center.

Most scholars of the movement are si-

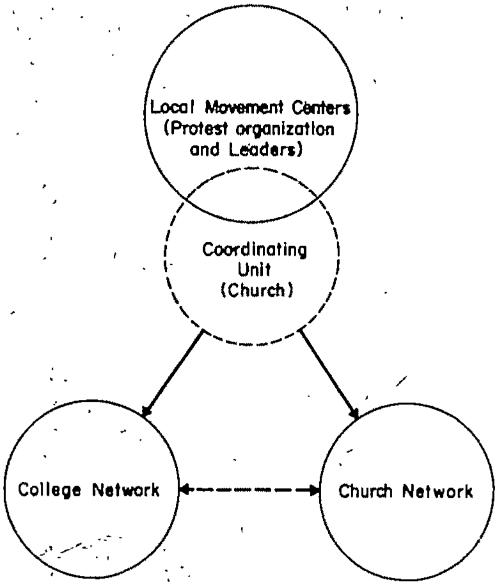


Figure 1. Structure of a Typical Local Movement Center

lent about the period between the Montgomery bus boycott and the 1960 sit-ins. My analysis emphasizes that the organizational foundation of the Civil Rights movement was built during this period and active local movement centers were created in numerous Southern black communities. For instance, between 1957 and 1960 many local centers emerged in Virginia. Ministers such as Reverends Milton Reid, L. C. Johnson, Virgil Wood, Curtis Harris, and Wyatt Walker operated out of centers in Hopewell, Lynchburg, Portsmouth, and Petersburg. The direct action organizations of these cities were named Improvement Associations and were patterned after the original MIA. South Carolina also had its movement centers. For example, in 1955–1956, after whites began exerting economic pressure against blacks desiring school integration, the black community of Orangeburg initiated an economic boycott against twenty-three local firms. This extended boycott resulted in a vibrant movement center led by the Reverends Matthew McCollom, William Sample, and Alfred Issac and their congregations. Movement centers emerged in other South Carolina cities, such as Sumter, Columbia, and Florence, organized by James McCain of CORE and activist clergymen.

In Durham, North Carolina, churches that made up the movement center were Union Baptist, pastored by Rev. Grady Davis; Ashbury Temple, pastored by Rev. Douglas Moore; Mount Zion, pastored by Rev. Fuller; St. Marks, pastored by Rev. Speaks; and St. Josephs, pastored by Rev. Swann. Movement centers were also to be found in cities of the deep South such as Montgomery and Birmingham, Alabama; Baton Rouge, Louisiana; and Tallahassee, Florida.

So prevalent were these centers throughout the South that when Gordon Carey, a CORE field investigator, surveyed the situation in 1959, he reported:

> In some Southern cities such as Montgomery, Orangeburg, Tallahassee, and Birmingham nonviolent movements have been and are being carried on. But most of the South, with its near total segregation, has not been touched. Many places have *felt* the *spirit* of Martin Luther King, Jr. but too often this spirit has not been turned into positive action. [Carey, 1959, emphasis added]

The "spirit" to which Carey referred was in fact the church-based movement centers he found throughout the South, most of which were affiliated with or patterned after SCLC.

Elsewhere (Morris, 1980), I have analyzed how, in the late 1950s, these centers were perfecting confrontation strategies, building organizations, leading marches, organizing voter drives, and radicalizing members of the community. Scholars (e.g., Oberschall, 1973:223) persistently dismiss these centers as weak, limited, and unwilling to confront the white power structure. Yet the evidence suggests a different interpretation. For example, Rev. Fred Shuttlesworth and his mass-based movement center continually confronted Bull Connor and the white power structure of Birmingham throughout the late fifties. As a consequence, Shuttlesworth's home and church were repeatedly bombed.

In short, between 1955 and 1960 many local movement centers were formed and hardened. These centers, which included NAACP Youth Councils and CORE chapters, constituted the new political reality of Southern black communities on the eve of the 1960 sit-ins. It was these

structures that were able to generate and sustain a heavy volume of collective action.

THE GREENSBORO CONNECTION

On February 1, 1960 Ezell Blair Jr., Franklin McCain, Joe McNeil, and David Richmond, all students at North Carolina Agricultural and Technical College, sat-in at the Woolworth's lunch counter in Greensboro, North Carolina. Though most commentators mark this as the first sit-in, the four protesters knew that they were not the first to sit-in in the state of North Carolina. Sit-in activity in the state had begun in the late fifties, when a young black attorney, Floyd McKissick, and a young Board member of SCLC, Rev. Douglas Moore, and a small group of other young people (including a few whites from Duke University) began conducting sit-ins in Durham.

These early Durham sit-ins were part of the network of sit-ins which occured between 1957 and 1960. The activists involved in the early sit-ins belonged to the NAACP Youth Division, which McKissick headed, and their own direct-action organization called the Durham Committee on Negro Affairs. During the late fifties, McKissick and Moore's group conducted sit-ins at local bus stations, waiting rooms, parks, hotels, and other places (McKissick, 1978). In 1957, Rev. Moore and others were arrested for sitting-in at a local ice-cream parlor. The subsequent legal case became known as the "Royal Ice Cream Case." McKissick, who also headed the local Boy Scout organization, periodically would take the young "all-American" scouts into segregated restaurants and order food. In short, this Durham group persistently confronted the white power structure in the late fifties.

The four students who sat-in at Greensboro and sparked the widespread sit-in movement had been members of the NAACP Youth Council, headed by McKissick. According to McKissick, he knew them all well and they knew all about the Durham activities. Martin Oppenheimer (1964:398), an early historian of the sit-ins, confirms this: "All of the boys were, or at some time had been members of an NAACP Youth Council." Indeed, the four students had participated in numerous meetings in social-action oriented churches in Durham. Involvement with the NAACP Youth Council meant that they were not only informed about the Durham sit-ins, but also knew about many of the sit-ins conducted prior to 1960. Thus, the myth that four college students got up one day and sat-in at Woolworth's—and sparked the movement—dries up like a "raisin in the sun" when confronted with the evidence.

The National office of the NAACP and many conservative ministers refused to back the Greensboro sit-ins. The NAACP's renowned team of lawyers did not defend the "Greensboro Four." Nevertheless, on the same day they sat-in, the students contacted a lawyer whom they considered to be their friend, and Floyd McKissick became the lawyer for the "Greensboro Four." The network of college students and adult activists had begun to operate in earnest.

Well-forged networks existed between and among black churches and colleges in North Carolina, facilitated by the large number of colleges concentrated in the state. Indeed, ten black colleges existed within a ten-mile radius of Greensboro (Wolff, 1970:590). Interactions between colleges and churches were both frequent and intense; many colleges were originally founded by the churches. A number of North Carolina churches were referred to as "college churches" because they had large student memberships. These two sets of social organizations were also linked through college seminaries where black ministers received their theological training.

These church-student networks enabled activist-oriented students to become familiar with the emerging Civil Rights movement via local movement centers and made it possible for adult activists to tap the organizational resources of the colleges. Leaders of student governments and other campus groups facilitated student mobilization because they, like the ministers, had organizing skills and access to blocs of people. Moreover, the concentration of colleges in the state provided an extensive network of contacts. Frater-

nity and sorority chapters linked students within and between campuses, as did dating patterns and joint cultural and athletic events. Finally, intercollegiate kinship and friendship networks were widespread, and student leaders were squarely tied to these networks. Similiarly, black communities across North Carolina could be rapidly mobilized through the churches, since churches were linked through ministerial alliances and other networks. By 1960 these diverse and interlocking networks were capable of being politicized and coordinated through existing movement centers, making North Carolina an ideal state for the rapid diffusion of collective action.

Within a week of the Greensboro protest, sit-ins rapidly spread across the South. In an extensive study, the Southern Regional Council (1960) reported that between February 1 and March 31 of 1960, major sit-in demonstrations and related activity had been conducted in at least sixty-nine Southern cities (see Table 1).[7]

Table 1. Number of Cities with Sit-ins and Related Protest Activities, February-March 1960, by State

State	Number
North Carolina	18
Florida	11
Virginia	9
South Carolina	7
Texas	5
Tennessee	4
Alabama	4
Georgia	2
West Virginia	2
Louisiana	2
Arkansas	2
Maryland	1
Ohio	1
Kentucky	1
Total	69

Compiled from: Southern Regional Council. "The student protest movement, winter 1960." SRC-13, April 1 1960 (revised)

BEYOND GREENSBORO

As soon as the sit-ins started in Greensboro, the network of movement centers was activated. In the first week of February, 1960, students continued to sit-in daily at the local Woolworth's, and the protest population began to grow. The original four protesters were joined by hundreds of students from A & T College and several other local black colleges. Black high-school students and a few white college students also joined the protest. Influential local whites decided to close the Woolworth's in Greensboro, hoping to take the steam out of the developing mass-movement. It was too late.

Floyd McKissick, Rev. Douglas Moore, and others who had conducted previous sit-ins formulated plans to spread the movement across the state. They were joined by CORE's white field secretary, Gordon Carey, whose services had been requested by the local NAACP president. Carey arrived in Durham from New York on February the 7th and went directly to McKissick's home, where the sit-ins were being planned. Carey was a good choice because he had knowledge of nonviolent resistance and because of his earlier contact with movement centers in Southern black communities.

On February 8th—exactly one week after the Greensboro sit-ins—the demonstrations spread to nearby Durham and Winston-Salem. McKissick, Moore, Carey, and others helped organize these sit-ins, bringing students from the local colleges to churches where they were trained to conduct sit-ins. For example, the Durham students were trained at the same churches through which McKissick and Moore had planned direct action in the late 1950s. Following training and strategy sessions, the students went to the local lunch counters and sat-in.

The organizing effort was not limited to these two nearby cities. Within the first week of the Greensboro sit-in, McKissick, Carey, and Rev. Moore made contact with activists in movement centers throughout North Carolina, South Carolina, and Virginia, urging them to train students for sit-ins. They not only phoned these activists, but traveled to various cities to provide assistance. Upon arrival they often found sit-in planning sessions al-

[7] To appreciate the volume of protest activity engendered by the sit-ins, it is necessary to note that the total number of cities (69) is not a count of actual day-to-day demonstrations, which during these first two months ran into the hundreds if not thousands.

ready underway. According to Carey (1978), "when we reached these cities we went directly to the movement oriented churches." When asked why, Carey replied, "Well, that's where the protest activities were being planned and organized." Thus, these sit-ins were largely organized at the movement churches rather than on the campuses. To understand the sit-in movement, one must abandon the assumption that it was a collegiate phenomenon. For different reasons, Rev. Moore attempted to convey this same idea in the early days of the sit-ins: "If Woolworth and other stores think this is just another panty raid, they haven't had their sociologists in the field recently" (Moore, 1960). The sit-ins grew out of a context of organized movement centers.

As anticipated above, the Southern Christian Leadership Conference was central to the rise of the 1960 sit-in movement. It is critical to remember that when Rev. Moore and other organizers visited churches in North and South Carolina and Virginia, they discovered that church leaders were already training students for sit-ins. Speaking of the ministers who headed these movement churches, Carey (1978) reported, "All of these ministers were active in the Southern Christian Leadership Conference. At least 75% were getting inspiration from King." Additionally, these ministers had contacts with and often were leaders of both CORE and the activist wing of the NAACP.

Since the movement centers were already in place, they served as both receiving and transmitting "antennas" for the sit-ins. As receivers they gathered information of the sit-ins, and as transmitters they rebroadcast information throughout the networks. Because this internal network already existed, information was rapidly channeled to groups prepared to engage in nonviolent collective action.

During the second week of February 1960, plans were formulated to conduct sit-ins in a number of Southern cities. Communication and coordination between the cities was intensified. For example, early in the second week of February, the Rev. B. Elton Cox of High Point, North Carolina, and Rev. C. A. Ivory of Rock Hill, South Carolina, phoned McKissick and other leaders, informing them that their groups were "ready to go" (McKissick, 1978). Cox's group sat-in on February 11th and Ivory's on February 12th. Rev. Ivory organized and directed the Rock Hill sit-ins from his wheelchair. Within the week, sit-ins were being conducted in several cities in Virginia, most of them organized through the dense network of SCLC movement centers in that state (Southern Regional Council, 1960; Walker, 1978).

The movement hot lines reached far beyond the border states of North Carolina, South Carolina, and Virginia. Rev. Fred Shuttlesworth, an active leader of the Birmingham, Alabama, movement center, happened to be in North Carolina when the first wave of sit-ins occurred, fulfilling a speaking engagement for the leader of the High Point sit-ins—Rev. Cox. According to Shuttlesworth, "He [Rev. Cox] carried me by where these people were going to sit-in . . . I called back to Atlanta, and told Ella [Baker] what was going on. I said, 'this is the thing. You must tell Martin [King] that we must get with this, and really this can shake up the world' " (Shuttlesworth, 1978). Baker, the Executive Director of SCLC, immediately began calling her contacts at various colleges, asking them, "What are you all going to do? It is time to move" (Baker, 1978).

Carey and Rev. Moore phoned the movement center in Nashville, Tennessee, and asked Rev. Lawson if they were ready to move. The student and church communities coordinated by the Nashville Christian Leadership Conference answered in the affirmative. According to Lawson,

> Of course there was organizing because after the sit-in, the first one in February, people like Doug Moore, Ella Baker, myself, did call around to places that we knew, said, 'Can you start? Are you ready? Can you go? And how can we help you?' So there was some of that too that went on. Even there the sit-in movement did not just spread spontaneously. I mean there was a readiness. And then there were, there were phone calls that went out to various communities where we knew people and where we knew student groups and where we knew minister groups,

> and said, you know, 'this is it, let's go.' [Lawson, 1978]

When asked, "Why did the student sit-in movement occur?" Lawson replied,

> Because King and the Montgomery boycott and the whole development of that leadership that clustered around King had emerged and was ready and was preaching and teaching direct action, nonviolent action, and was clearly ready to act, ready to seed any movement that needed sustenance and growth. So there was . . . in other words, the soil had been prepared. [Lawson, 1978]

These data provide insight into how a political movement can rapidly spread between geographically distant communities. The sit-ins spread across the South in a short period of time because activists, working through local movement centers, planned, coordinated, and sustained them. They spread despite the swinging billy clubs of policemen, despite Ku Klux Klansmen, white mobs, murderers, tear gas, and economic reprisals (Southern Regional Council, 1960; Matthews and Prothro, 1966; Oberschall, 1973). The pre-existing movement centers provided the resources and organization required to sustain the sit-ins in the face of opposition.

SIT-IN CLUSTERS OF 1960

The organizational and personal networks that produced the first cluster of sit-ins in Oklahoma in 1958 have already been described. The cluster concept can be applied to the entire set of sit-ins of February and March 1960. Many of the cities where sit-ins occurred can be grouped by geographic and temporal proximity. A cluster is defined as two or more cities within 75 miles of each other where sit-in activity took place within a span of 14 days. In Table 2, forty-one of the sixty-nine cities having sit-ins during this two-month period have been grouped because they meet these criteria. Within this period 59% of the cities that had sit-ins and related activity were part of clusters. The percentage of these cities forming sit-in clusters is even more striking in the first month: during February, 76% of cities having sit-ins were part of clusters, while during March the percentage dropped to 44%.

The clustering differentials between the two months can be explained by taking region into account as shown in Table 3. In the first month (February) 85% of the cities having sit-ins were located in Southeastern and border states. This pattern had been established earlier, when most of the pre-1960 sit-ins occurred in border states. Most of the February sit-ins took place in cities of border states because repression against blacks was not as severe there as in the deep South. This made it possible for activists in border states to build dense networks of movement centers. We have already seen that North Carolina, South Carolina, and Virginia had numerous social-action churches and direct-action organizations. By the time the sit-ins occurred in Virginia, SCLC had affiliates throughout the state, and Rev. Wyatt Walker, who was the leader of Virginia's movement centers, was also the state Director of CORE and President of the local NAACP. Similar patterns existed in the other border states. Small wonder that in the month of February, 73% of cities having sit-ins were located in Virginia, North Carolina, and South Carolina. Similarly, these cities produced 88% of the February clusters. This clustering reflected both the great density of movement centers and a system of domination less stringent than that of the deep South.

Table 3 reveals that in March a major change took place: the majority of the sit-ins occurred in cities of the deep South. With a few exceptions, the sit-ins in the deep South did not occur in clusters. They occurred almost exclusively in Southern cities where movement centers were already established: Montgomery and Birmingham, Alabama; Baton Rouge and New Orleans, Louisiana; Tallahassee, Florida; Nashville and Memphis, Tennessee; and Atlanta and Savannah, Georgia. Repression would have been too great on student protesters operating outside of the protection of such centers in the deep South. Thus, the decrease in clustering in the deep South reflected both the high level of repression and the absence of dense networks of movement centers.

Table 2. Clusters of Cities with Sit-ins and Related Activities, February–March 1960

Cluster	Number of days between first sit-ins within cluster	Maximum number of miles between farthest two cities within cluster
Fayetteville, Raleigh, N.C. (2/9/60–2/10/60)	1	50
Tampa, St. Petersburg, Sarasota, Fla. (2/29/60–3/2/60)	2	50
Montgomery, Tuskegee, Ala. (2/25/60–2/27/60)	2	25
Columbia, Florence, Sumter, S.C. (3/2/60–3/4/60)	2	70
Austin, San Antonio, Tx. (3/11/60–3/13/60)	2	75
Salisbury, Shelby, N.C. (2/16/60–2/18/60)	2	60
Wilmington, New Bern, N.C. (3/17/60–3/19/60)	2	75
Charlotte, N.C., Concord, Rock Hill, S.C. (2/9/60–2/12/60)	3	50
Durham, Winston-Salem, High Point, N.C. (2/8/60–2/11/60)	3	75
Chapel Hill, Henderson, N.C. (2/25/60–2/28/60)	3	50
Jacksonville, St. Augustine, Fla. (3/12/60–3/15/60)	3	40
Charleston, Orangeburg, Denmark, S.C. (2/25/60–2/29/60)	4	70
Daytona Beach, Sanford, Orlando, Fla. (3/2/60–3/7/60)	5	54
Houston, Galveston, Tx. (3/5/60–3/11/60)	6	65
Richmond, Petersburg, Va. (2/20/60–2/27/60)	7	30
Hampton, Norfolk, Portsmouth, Suffolk, Newport News (2/11/60–2/22/60)	11	35

Compiled from: Southern Regional Council. "The student protest movement, winter 1960." SRC-13, April 1 1960

Focusing on the internal movement centers enables us to explain both the clustering phenomenon and its absence.

Given the large proportion of sit-ins occurring in clusters, we can say that they did not spread randomly. The clusters represented the social and temporal space in which sit-ins were organized, coordinated, spread, and financed by the black community.[8] Within these clusters, cars filled with organizers from SCLC, NAACP, and CORE raced between sit-in points relaying valuable information. Telephone lines and the community "grapevine" sent forth protest instructions and plans. These clusters were the sites of numerous midday and late night meetings where the black community assembled in the churches, filled the collection plates, and vowed to mortgage their homes to raise the necessary bail-bond money in case the protesting students were jailed. Black lawyers pledged their legal services to the movement and black physicians made their services available to injured demonstrators. Amidst these exciting scenes, black spirituals that had grown out of slavery calmed and deepened the participants' commitment. A detailed view of the Nashville sit-ins provides an example of these dynamics, because the Nashville movement epitomized the sit-ins whether they occurred singularly or in clusters.

[8] Cities identified as part of a particular cluster may actually be part of another cluster(s). I assume that the probability of shared organization and coordination of sit-ins is high if two or more cities within a 75-mile radius had sit-ins within a two-week period. My data and analysis generally confirm this assumption.

Table. 3. Cities with Sit-ins and Related Activities, February–March 1960, by Geographic Region

	Deep South	Southeastern and Border States	Non-South	All States
February–March 1960				
Number of cities with sit-ins, 2-month total	26	42	1	69
Region's % of 2-month total	38	61	1	100
February 1960				
Number of cities with sit-ins	5	28	0	33
Region's % of Feb. total	15	85	0	100
% of 2-month total occurring in Feb.	19	67	0	48
March 1960				
Number of cities with sit-ins	21	14	1	36
Region's % of March total	58	39	3	100
% of 2-month total occurring in March	81	33	100	52

Compiled from: Southern Regional Council. "The student protest movement, winter 1960." SRC-13, April 1 1960

NOTE: Deep South states are Alabama, Florida, Georgia, Texas, Arkansas, and Louisiana. Southeastern and Border states are South Carolina, North Carolina, Virginia, Tennessee, Maryland, Kentucky, and West Virginia. The non-South state is Ohio.

THE NASHVILLE SIT-IN MOVEMENT

A well-developed, church-based movement center headed by Rev. Kelly Miller Smith was organized in Nashville during the late 1950s. The center, an affiliate of SCLC, was called the Nashville Christian Leadership Council (NCLC). Rev. James Lawson, an expert tactician of nonviolent protest, was in charge of NCLC's direct-action committee. Lawson received a call from Rev. Douglas Moore about two days after the Greensboro sit-ins began. The Nashville group was ready to act because a cadre of students had already received training in nonviolent direct action. They had conducted "test sit-ins" in two large department stores in downtown Nashville prior to the 1959 Christmas holidays. Moreover, the group had already made plans in late 1959 to begin continuous sit-ins in 1960 with the explicit intention of desegregating Nashville (Smith, 1978; D. Bevel, 1978). Thus, Greensboro provided the impetus for the Nashville group to carry out its pre-existing strategy.

Rev. Smith's First Baptist Church became the coordinating unit of the Nashville sit-in movement. A decision to sit-in at local lunch counters on Saturday, February 13 1960, was arrived at after much debate. The adults (mostly ministers) of the NCLC met with the students at movement headquarters and tried to convince them to postpone the demonstrations for a couple of days until money could be raised. According to Rev. Smith (1978), "NCLC had $87.50 in the treasury. We had no lawyers, and we felt kind of a parental responsibility for those college kids. And we knew they were gonna be put in jail, and we didn't know what else would happen. And so some of us said, 'we need to wait until we get a lawyer, until we raise some funds.' "

NCLC leaders told the students that they could collect the money through the churches within a week. Then, according to Rev. Smith:

> James Bevel, then a student at American Baptist Theological Seminary, said that, 'I'm sick and tired of waiting,' which was a strange thing to come from a kid who was only about nineteen years old. You see, the rest of us were older . . . [Bevel said] 'If you asked us to wait until next week, then next week something would come up and you'd say wait until the next week and maybe we never will get our freedom.' He said this, 'I believe that something will happen in the situation that will make for the solution to some of these problems we're talking about.' So we decided to go on. [Smith, 1978]

The proximity of four black colleges in Nashville—Fisk University, Tennessee State College, American Baptist Theological Seminary, and Meharry Medical School—facilitated the mobilization of large numbers of students. In its

extensive ties between students and churches, Nashville resembled the state of North Carolina. Indeed, John Lewis, James Bevel, and Bernard Lafayette, who became major sit-in leaders, were students at the American Baptist Theological Seminary and were taught there by Rev. Smith. Furthermore, they were student leaders:

> John Lewis, Bernard and myself were the major participants in the seminary. All of us were like the top student leaders in our schools. I think John at the time was the president of the Student Council. I was a member of the Student Council. I was one of the editors of the yearbook. Bernard was an editor of the yearbook. So all of us were like the top leaders in our school. [J. Bevel, 1978]

Thus the student leaders could rapidly mobilize other students because they already had access to organized groups. Other writers (Von Eschen et al., 1971; McAdam, 1979) have pointed out that these college networks played a key role in sit-in mobilization. However, the sit-in movement cannot be explained without also noting the crucial interaction between black college students and local movement centers. Speaking of Rev. Smith and his church, Bevel recalled, "the First Baptist basically had the Baptist people who went to Fisk and Meharry and Tennessee State, and the Seminary were basically members of his church" (J. Bevel, 1978). These students had been introduced to the Civil Rights movement while they attended church.

On the first day of the sit-ins in Nashville, students gathered in front of their respective campuses. NCLC sent cars to each college to transport the students to Rev. Smith's church. Again, the major organizational tasks were performed in the church which served as the coordinating unit of the local movement center, rather than on the campuses. Coordination of sit-in activity between the college community and the churches was made less difficult because many of the students (especially student leaders) were immersed in the local movement centers prior to the sit-ins. The pattern of close connection between student demonstrators and adult leaders had already existed in places such as Greensboro and even Oklahoma City in 1958; indeed, this pattern undergirded the entire movement. Rev. Jemison's (1978) remark that the Baton Rouge sit-in demonstrators "were schooled right over there at our church; they were sent out from here to go to the lunch counters" typifies the relationship between the students and the local movement centers.[9] Jemison continued, "The student leaders attended church here. We had close ties because they were worshipping with us while we were working together."

Once the Nashville students arrived at movement headquarters, they participated in workshops where they learned the strategies of nonviolent confrontation from experts like Rev. Lawson, Rev. Metz Rollins, Rev. C. T. Vivian, and the core group of students that Lawson had already trained. This pool of trained leaders was a pre-existing resource housed by NCLC. After the workshops, the students were organized into groups with specific protest responsibilities, each having a spokesperson who had been trained by Lawson during the late 1950s. They then marched off to confront Nashville's segregated lunch counters and agents of social control.

The adult black community immediately mobilized to support the students. Shortly after the demonstrations began, large numbers of students were arrested. According to Rev. Smith,

> We just launched out on something that looked perfectly crazy and scores of people were being arrested, and paddy wagons were full and the people out in downtown couldn't understand what was going on, people just welcoming being arrested, that ran against everything they had ever seen. . . . I've forgotten how much we needed that day, and we got everything we needed. [That particular day?] Yes, sir. About $40,000. We needed something like $40,000 in fives. And we had all the money. Not in fives, but in bail. Every bit of it came up. You know—property and this kind of thing . . . and there were fourteen black lawyers in this town. Every black lawyer made himself available to us. [Smith, 1978]

---

[9] For further evidence of the centrality of student-church ties in other cities that had sit-ins see Morris, forthcoming.

Thus, basic, pre-existing resources in the dominated community were used to accomplish political goals. It was suggested to Rev. Smith that a massive movement such as that in Nashville would need outside resources. He replied,

> Now let me quickly say to you that in early 1960, when we were really out there on the line, the community stood up. We stood together. This community had proven that this stereotyped notion of black folk can't work together is just false. We worked together a lot better than the white organizations. So those people fell in line. [Smith, 1978]

Rev. Smith's comments are applicable beyond Nashville. For example, in Orangeburg, after hundreds of students were arrested and brutalized, the adult black community came solidly to their aid. Bond was set at $200 per student, and 388 students were arrested. Over $75,000 was needed, and adults came forth to put up their homes and property in order to get students out of jail. Rev. McCollom, the leader of the Orangeburg movement center, remarked that, "there was no schism between the student community and the adult community in Orangeburg" (McCollom, 1978). Jim McCain (1978) of CORE, who played a central role in organizing sit-ins across South Carolina and in Florida, reported that community support was widespread. According to Julian Bond (1980), a student leader of Atlanta's sit-ins, "black property owners put up bond which probably amounted to $100,000" to get sit-in demonstrators released from jail.

These patterns were repeated across the South. This community support should not be surprising, considering the number of ministers and congregations involved before and during the movement. Yet, Zinn, an eyewitness to many of these events, wrote, "Spontaneity and self-sufficiency were the hallmarks of the sit-ins; without adult advice or consent, the students planned and carried them through" (1964:29). This myopia illustrates the inadequacies of analyses that neglect or ignore the internal structure of oppressed communities and protest movements.

The continuing development of the Nashville sit-ins sheds further light on the interdependence of the movement and the black community. A formal structure called the Nashville Nonviolent Movement was developed to direct sit-in activities. Its two substructures, the Student Central Committee and the Nashville Christian Leadership Council, worked closely together and had overlapping membership (Reverends Lawson and Vivian were members of both groups). The Central Committee usually consisted of 25 to 30 students drawn from all the local colleges. NCLC represented adult ministers and the black community. The two groups established committees to accomplish specific tasks, including a finance committee, a telephone, publicity, and news committee, and a work committee. The work committee had subgroups responsible for painting protest signs and providing food and transportation. The city's black lawyers became the movement's defense team, students from Meharry Medical School were the medical team.

This intricate structure propelled and guided the sit-in movement of Nashville. A clear-cut division of labor developed between the Central Committee and the NCLC. The Central Committee's major responsibilities were to train, organize, and coordinate the demonstration. The NCLC developed the movement's financial structure and coordinated relations between the community and the student movement. Diane Nash Bevel, a major student leader of the Nashville sit-ins, was asked why the students did not take care of their own finances and build their own relationships with the larger community. She replied,

> We didn't want to be bothered keeping track of money that was collected at the rallies and stuff. We were just pleased that NCLC would do that, and would handle the bookkeeping and all that trouble that went along with having money. . . . Besides, we were much too busy sitting-in and going to jail and that kind of thing. There wasn't really the stability of a bookkeeper, for instance. We didn't want to be bothered with developing that kind of stability. . . . We were very pleased to form this alliance with NCLC who would sponsor the rallies and coordinate the community support among the adults and

> keep track of the money, while we sat-in and . . . well, it took all our time, and we were really totally immersed in it. My day would sometimes start . . . well we'd have meetings in the morning at six o'clock, before classes, and work steady to extremely late at night, organizing the sit-ins, getting publicity out to the students that we were having a sit-in, and where and what time we would meet. Convincing people, and talking to people, calming people's fears, going to class, at the same time. It was a really busy, busy time for all of the people on the Central Committee. We were trying to teach nonviolence, maintain order among a large, large number of people. That was about all we could handle. [D. Bevel, 1978]

Students are ideal participants in protest activities. Usually they do not have families to support, employer's rules and dictates to follow, and crystallized ideas as to what is "impossible" and "unrealistic." Students have free time and boundless energy to pursue causes they consider worthwhile and imperative (Lipset and Wolin, 1965:3; McCarthy and Zald, 1973:10). McPhail's (1971:1069) finding that young, single, unemployed males were ideal participants in civil disorders and McPhail and Miller's (1973:726) discussion of availability for participation in the assembly process parallels this notion that students are ideal participants in protest activities. Nevertheless, although black students were able to engage in protest activities continuously because of their student status, a one-sided focus on them diverts attention from the larger community, which had undergone considerable radicalization. Speaking of the adults, James Bevel (1978), a student organizer of the Nashville sit-ins, remarked, "But when you talk to each individual, they talked just like we talked—the students. They had jobs and they were adults. But basically, their position would be just like ours. They played different roles because they were in different—they had to relate based on where they were in the community" (J. Bevel, 1978).

The adults of the NCLC organized the black community to support the militant student sit-in movement. Once the movement began, NCLC instituted weekly and sometimes daily mass meetings in the churches. Rev. Smith (1978) recalled,

> Sometimes we had them more than once a week if we needed to. When things were really hot we called a meeting at eight o'clock in the morning. We'd call one for twelve that day, twelve noon, and the place would be full. We had what we called our wire service. People got on telephones, that was our wire service, and they would fill that building. They'd fill that building in just a matter of relatively short time."

At these mass meetings, ministers from across the city turned over the money that their respective churches had donated to the movement. Thousands of dollars were collected at the mass meetings while black adults, ministers, and students sang such lyrics as "Before I'd be a slave, I'd rather be buried in my grave." Then too, bundles of leaflets were given to adults at mass meetings who then distributed them throughout the black community. This shows how the movement built communication channels through which vital information, strategies, and plans were disseminated.

During the Nashville sit-ins, word went out to the black community not to shop downtown.

> We didn't organize the boycott. We did not organize the boycott. The boycott came about. We don't know how it happened. I tell you there are a lot of little mystical elements in there, little spots that defy rational explanation. . . . Now, we promoted it. We adopted it. But we did not sit down one day and organize a boycott . . . ninety-nine percent of the black people in this community stayed away from downtown during the boycott. It was a fantastic thing—successful. It was fantastically successful. [Smith, 1978]

Yet the boycott was largely organized by NCLC. According to Bevel, Dr. Vivian Henderson, who was head of Fisk University's economic department and a member of NCLC, played a key role in the boycott, because

> Vivian Henderson was basically responsible for calling the boycott. He got up at a mass meeting and said, 'at least what we could do to support students, if we've got any decency, we can just stop paying bills and just don't shop until this thing is resolved.' A very indignant type of speech he made. It just caught on. All the bourgeois women would come to the meeting, and they just got on the phone and called up everybody, all

> the doctors' wives and things. They just got on the phone and called 300 or 400 people and told them don't shop downtown. Finally there was just a total boycott downtown. There would be no black people downtown at all. [J. Bevel, 1978]

Activists were stationed downtown to insure that blacks knew not to shop. According to Rev. Smith, shortly after the boycott was initiated, merchants began coming to his home wanting to talk. Diane Nash Bevel attributed the boycott's effectiveness to reduced profits during the Easter shopping season. It also changed the merchant's attitude toward the sit-ins.

> It was interesting the difference that [the boycott] made in terms of how the managers were willing to talk with us, because see we had talked with the managers of the stores. We had a meeting at the very beginning and they had kind of listened to us politely, and said, 'well, we just can't do it. We can't desegregate the counters because we will lose money and that's the end of it.' So, after the economic withdrawal, they were eager to talk with us, and try to work up some solution. [D. Bevel, 1978]

In early 1960 the white power structure of Nashville was forced to desegregate a number of private establishments and public transportation facilities. SNCC's *Student Voice* reported that in Nashville, "A long series of negotiations followed the demonstrations, and on May 10, 6 downtown stores integrated their lunch counters. Since this time others have followed suit, and some stores have hired Negroes in positions other than those of menial workers for the first time" (*Student Voice*, August, 1960). Daily demonstrations by hundred of students refusing to accept bond so that they could be released from jail, coupled with the boycott, gave blacks the upper hand in the conflict situation. Careful organization and planning was the hallmark of the Nashville sit-in movement.

DISCUSSION AND CONCLUSIONS

Consistent with Proposition 1, I have presented evidence that pre-existing social strctures played a central role in the 1960 sit-in movement. Pre-existing activist groups, formal movement organizations, colleges, and overlapping personal networks provided the framework through which the sit-ins emerged and spread. Previous writings on the sit-ins (e.g., Lomax, 1962; Zinn, 1964; Matthews and Prothro, 1966; Killian, 1968; Meier and Rudwick, 1973; Piven and Cloward, 1977) have persistently portrayed pre-existing organization as an after-the-fact accretion on student spontaneity. The dominant view is that SCLC, CORE, NAACP, and community leaders rushed into a dynamic campus movement after it was well underway, while my data provide evidence that those organizational and community forces were at the core of the sit-in movement from its beginning. Thus, pre-existing organizations provided the sit-ins with the resources and communication networks needed for their emergence and development.

Prior to 1960 the sit-in was far from being the dominant tactic of the Civil Rights movement, yet in early 1960, sit-in demonstrations swept through thirteen states and hundreds of communities within two months. Almost instantly sit-ins became the major tactic and focus of the movement. A tactical innovation had occurred.

Consistent with Proposition 2, the data strongly suggest that the 1960 Greensboro sit-in occurred at the time when the necessary and sufficient condition for the rapid diffusion of sit-ins was present. That condition was the existence of well-developed and widespread internal organization. Because this internal organization was already firmly in place prior to 1960, activist groups across the South were in a position to quickly initiate sit-ins. The rapidity with which sit-ins were organized gave the appearance that they were spontaneous. This appearance was accentuated because most demonstrators were students rather than veteran Civil Rights activists.

Yet the data show that the student organizers of the sit-ins were closely tied to the internal organization of the emerging Civil Rights movement. Prior student/activist ties had been formed through church affiliations and youth wings of Civil Rights organizations. In short, students and seasoned activists were able to rapidly coordinate the sit-ins because both

were anchored to the same organization.

Innovations in political movements arise in the context of an active opposition. The organization of the Civil Rights movement provided the resources that sustained diffusion of the sit-ins in the face of attack. This vast internal organization consisted of local movement centers, experienced activists who had amassed organizing skills, direct-action organizations, communication systems between centers, pre-existing strategies for dealing with the opposition, workshops and training procedures, fund-raising techniques, and community mobilization techniques.

The pre-existing internal organization enabled organizers to quickly disseminate the "sit-in" idea to groups already favorably disposed toward direct action. In the innovation/diffusion literature (e.g., Coleman et al., 1957; Lionberger, 1960; Rogers, 1962) a positive decision by numerous actors to adopt a new item is treated as a central problem. In the case of the sit-ins, the adoption problem was largely solved by the pre-existing organization. Since that organization housed groups that had already identified with "confrontational politics," little time was lost on debates as to whether sit-ins should be adopted. Thus, the diffusion process did not become bogged down at the adoption stage.

Repression might have prevented the diffusion process. The authorities and white extremist groups attempted to prevent the spread of the sit-ins by immediately arresting the demonstrators, employing brutal force, and refusing to report the sit-ins in the local press. The organizational efficiency of the movement centers prevailed against the opposition. Existing recruiting and training procedures made it possible for jailed demonstrators to be instantly replaced. When heavy fines were leveled against the movement, activists were able generally to raise large sums of money through their pre-existing community contacts. The pre-existing communication networks easily overcame the problems imposed by news blackouts. Moreover, skilled activists were able to weaken the stance of the opposition by rapidly organizing economic boycotts. Because the internal organization was widespread, these effective counter measures were employed in Black communities across the South. Thus, it was well-developed and widespread internal organization that enabled the 1960 sit-ins to rapidly diffuse into a major tactical innovation of the Civil Rights movement.

Propositon 3 maintains that pre-existing internal organization establishes the types of innovations that can occur within movements. The internal organization that gave rise to the sit-ins specialized in what was called nonviolent direct action. This approach consisted of a battery of tactics that were disruptive but peaceful. The nonviolent approach readily fitted into the ideological and organizational framework of the black church, and provided ministers, students, and ordinary working people with a method for entering directly into the political process.

The movement centers that emerged following the Montgomery bus boycott were developed around nonviolent approaches to social change. Indeed, the primary goal of these centers was to build nonviolent movements. Yet, nonviolent confrontations as a disciplined form of collective action was relatively new to the black masses of the South. The activists within the movement centers systematically introduced blacks to the nonviolent approach. They organized nonviolent workshops and conducted them on a routine basis in the churches and protest organizations. Literature from organizations (e.g., Fellowship of Reconciliation and CORE) that specialized in the nonviolent approach was made available through the centers. Skilled nonviolent strategists (e.g., Bayard Rustin, James Lawson, and Glenn Smiley) travelled between centers training leaders how to conduct nonviolent campaigns. The varied tactics—mass marches, negotiations, boycotts, sit-ins—associated with direct action became common knowledge to activists in the centers. Moreover, in the late fifties activists began experimenting with these tactics and urging the community to become involved with nonviolent confrontations. Meier and Rudwick (1976) have shown that sit-ins at segregated

facilities were conducted by black activists in the nineteen forties and late fifties. But this tactic remained relatively isolated and sporadic and did not diffuse throughout the larger community. Meier and Rudwick (1976:384) conclude that diffusion did not occur before 1960 because the white mass-media failed to cover sit-ins. My analysis suggests another explanation: sit-ins prior to 1960 did not spread because the internal organization required for such a spread did not exist. In short, without viable internal social organization, innovations will remain sporadic and isolated. With organization, innovations can spread and be sustained. By 1960 the internal organization of the Civil Rights movement had amassed resources and organization specifically designed to execute nonviolent confrontations.

The sit-in tactic was well suited to the existing internal organization of the Civil Rights movement. It did not conflict with the procedures, ideology, or resources of the movement centers. Indeed, because the sit-in method was a legitimate tactic of the direct-action approach, it was quickly embraced by activists situated in the movement centers. Because these activists were already attempting to build nonviolent movements, they instantly realized that massive sit-ins could have a wide impact. Furthermore, they were well aware that they were in command of precisely the kinds of resources through which the sit-ins could be rapidly diffused. This is why they phoned activist groups and said, "This is it, let's go!" That is, the sit-ins became a tactical innovation within the movement because they fit into the framework of the existing internal organization.

In conclusion, this paper has attempted to demonstrate the important role that internal organization played in the sit-in movement. It is becoming commonplace for writers (e.g., Hubbard, 1968; Lipsky, 1968; Marx and Useem, 1971; McCarthy and Zald, 1973; Oberschall, 1973) to assert that the Civil Rights movement was dependent on outside resources: elites, courts, Northern white liberals, mass media, and the Federal Government. The present analysis suggests that this assertion may be premature, especially when the role of internal organization is ignored. Future research on collective action that treats internal organization as a topic in its own right will further increase our knowledge of the dynamics of social movements.

REFERENCES

Baker, Ella
1978 Interview. New York, New York.

Bevel, Diane Nash
1978 Interview. Chicago, Illinois. December 14.

Bevel, James
1978 Interview. New York, New York. December 27.

Blumer, H.
1946 "Collective Behavior." Pp. 165–220 in A. M. Lee (ed.), New Outline of the Principles of Sociology. New York: Barnes & Noble.

Bond, Julian
1980 Interview. Ann Arbor, Michigan. October 19.

Carey, Gordon
1959 Report to CORE National Council. February 21–22.
1978 Interview. Soul City, North Carolina. November 18 (Follow-up telephone interview November 1, 1979).

Coleman, James S., Eliher Katz, and Herbert Menzel
1957 "The diffusion of an innovation among physicians." Sociometry 20:253–70.

Frazier, E. Franklin
1963 The Negro Church in America. New York: Schocken Books.

Gamson, William A.
1975 The Strategy of Social Protest. Homewood, Illinois: Dorsey Press.

Haber, Robert A.
1966 "From protest to radicalism: an appraisal of the student struggle 1960." Pp. 41–9 in Mitchell Cohen and Dennis Hale (eds.), The New Student Left. Boston: Dorsey.

Hubbard, Howard
1968 "Five long hot summers and how they grew." Public Interest 12:3–24.

Jemison, Rev. T. J.
1978 Interview. Baton Rouge, Louisiana. October 16.

Jenkins, J. Craig and Charles Perrow
1977 "Insurgency of the powerless: farm workers movements (1946–1972)." American Sociological Review 42:249–68.

Killian, Lewis M.
1968 The Impossible Revolution? New York: Random House.

Lang, Kurt and Gladys Lang
1961 Collective Dynamics. New York: Crowell.

Lawson, James
1978 Interview. Los Angeles, California. October 2 and 6.

Lewis, John
1978 Interview. Washington, D.C. November 9.

Lewis, Chester
1981 Interview. Wichita, Kansas. February 3.

Lionberger, H. F.
1960 Adoption of New Ideas and Practices. Ames: The Iowa State University Press.
Lipset, Seymour Martin and Sheldon S. Wolin
1965 The Berkeley Student Revolt. Garden City, New York: Doubleday.
Lipsky, Michael
1968 "Protest as a political resource." American Political Science Review 62:1114–58.
Lomax, Louis E.
1962 The Negro Revolt. New York: New American Library.
Luper, Clara
1979 Behold the Walls. Jim Wire.
1980 Interview. Oklahoma City, Oklahoma. (Follow-up interview, January 1981).
Marx, Gary T. and Michael Useem
1971 "Majority involvement in minority movements: civil rights, abolition, untouchability." Journal of Social Issues 27:81–104.
Mays, Benjamin and Joseph W. Nicholson
1933 The Negro's Church. New York: Arno Press and the New York Times.
Matthews, Donald and James Prothro
1966 Negroes and the New Southern Politics. New York: Harcourt, Brace, and World.
McAdam, Douglas
1979 "Political process and the civil rights movement 1948–1962." Ph.D. dissertation, Department of Sociology, State University of New York at Stony Brook.
McCain, James
1978 Interview. Sumter, South Carolina. November 18.
McCarthy, J. D. and M. N. Zald
1973 The Social Trends of Social Movements in America: Professionalism and Resource Mobilization. Morristown, N.J.: General Learning Press.
McCollom, Rev. Matthew
1979 Interview. Orangeburg, South Carolina. October 31.
McKissick, Floyd
1978 Interview. Soul City, North Carolina. November 18 (Follow-up telephone interview November 2, 1979).
McPhail, Clark
1971 "Civil disorder participation; a critical examination of recent research." American Sociological Review 36:1058–73.
1973 "The assembling process: a theoretical and empirical examination." American Sociological Review 38:721–35.
Meier, August and Elliot Rudwick
1966 From Plantation to Ghetto. New York: Hill and Wang.
1973 CORE A study in the Civil Rights Movement 1942–1968. Oxford University Press.
1976 Along the Color Line. University of Illinois Press.
Moore, Douglas
1960 Journal and Guide. Vol. LX, March 5, 1960.
1978 Interview. Washington, D.C. November 1.
Morris, Aldon
1980 "The origins of the civil rights movement: an indigenous perspective." Ph.D. dissertation, Department of Sociology, State University of New York at Stony Brook.
forthcoming Origins of the Civil Rights Movement. New York: Free Press.
Obear, Frederick W.
1970 "Student activism in the sixties." Pp. 11–26 in Julian Foster and Durward Long (eds.), Protest: Student Activism in America 1970. New York: William Morrow.
Oberschall, Anthony
1973 Social Conflict and Social Movements. Englewood Cliffs, N.J.: Prentice Hall.
Oppenheimer, Martin
1964 "The southern student movement: year I." Journal of Negro Education 33:396–403.
Piven, Frances Fox, and Richard A. Cloward
1977 Poor People's Movements. New York: Vintage
Randolph, Homer
1981 Interview. East St. Louis, Illinois.
Rogers, Everett M.
1962 Diffusion of Innovations. New York: The Free Press of Glencoe.
Sale, Kirkpatrick
1973 SDS. New York: Vintage.
Shuttlesworth, Rev. Fred
1978 Interview. Cincinnati, Ohio. September 12.
Smelser, Neil J.
1963 Theory of Collective Behavior. New York: Free Press.
Smith, Rev. Kelly Miller
1978 Interview. Nashville, Tennessee. October 13.
Southern Regional Council
1960 "The student protest movement, winter 1960." SRC-13, April 1.
Student Nonviolent Coordinating Committee
1960 The Student Voice. August.
Tilly, Charles
1978 From Mobilization to Revolution. Reading, Massachusetts: Addison-Wesley.
Tilly, Charles, Louise Tilly, and Richard Tilly
1975 The Rebellious Century 1830–1930. Cambridge, Massachusetts: Harvard University Press.
Turner, Ralph and Lewis Killian
1957 Collective Behavior. Englewood Cliffs, N.J.: Prentice-Hall.
Von Eschen, Donald, Jerome Kirck, and Maurice Pinard
1971 "The organizational substructure of disorderly politics." Social Forces 49:529–44.
Walker, Rev. Waytt Tee
1978 Interview. New York City, New York. September 29.
Westby, David L.
1976 The Clouded Vision: The Student Movement in the United States in the 1960s. London: Associated University Press.
Williams, Hosea
1978 Interview. Atlanta, Georgia. September 22.
Wolff, Miles
1970 Lunch at the Five and Ten. New York: Stien and Day.
Zinn, Howard
1964 SNCC: The New Abolitionists. Boston: Beacon.

INEQUALITY WITHIN RURAL COMMUNITIES OF INDIA*

JOHN W. GARTRELL
University of Alberta

One of the principal controversies in development research has revolved around questions of whether and how the production of greater economic surplus in society results in greater or lesser inequality. Recent attempts to examine these issues within a world-systems context have employed secondary analysis of cross-national surveys of income inequality (Chase-Dunn, 1975; Rubinson, 1976; Bornschier et al., 1978; Weede, 1980; Stack, 1978). These efforts have had to rely upon aggregate data of sometimes dubious reliability and questionable theoretical appropriateness.

The present research examines uneven development within a single state (Andhra Pradesh) in one underdeveloped nation (India). It studies 84 agrarian villages, using the rural community as the unit of analysis. The study attempts to generate valid data through the use of both archival methods and first-hand village-level surveys.

COMMUNITY DEVELOPMENT AND DEPENDENCE

Unlike modernization theory, which describes traditional and modern sectors as separate economies, dependency theory describes linkages, for instance, between the village in rural India and the larger society. Dependency theory explains national and international inequality as inherent to the world capitalist system (Frank, 1967; Dos Santos, 1970; Wallerstein, 1974). The relationship between village communities and the wider society is viewed through the metaphor of internal colonialism (Portes, 1976; Portes and Walton, 1976; Griffin, 1978; Hechter, 1975; Davis, 1971); economic inequality results from uneven development and the dependence of the hinterland on the metropolis (Frank, 1966; Amin, 1976).

Looking from the bottom up, uneven development within underdeveloped nations (and developed ones, for that matter) can be seen as variation across communities. Community development can be measured in terms of the extent of integration into national (and international) production and marketing systems. Integration with national systems increases the community's dependence upon the wider society (Snipp and Summers, 1980), which in turn produces higher levels of inequality within the community (Frank, 1966; Chilcotte and Edelstein, 1974).

Yet modernization theory, which predicts that development should reduce inequality, cannot be lightly dismissed. Dalton has described village development and dependence as two sides of the same coin:

> Sustained income growth for the village community requires enlarged production for sale to regional, national, or international markets, and a return flow of consumption goods, producers goods, and social services (health and education) purchased with the ever-increasing cash income. . . . Agricultural production for sale (and) the use of modern technology . . . involve activities carried on with persons outside the local village, and create what might be called new dependency relationships, and gov-

* Direct all correspondence to Department of Sociology, University of Alberta, Edmonton, Alberta, T6G 2H4.

The field work upon which this report is based was supported by the Shastri Indo-Canadian Institute, The Foreign Area Fellowship Program, and the Midwestern Universities Consortium for International Activities. Writing and analysis were supported by a grant from the Social Sciences and Humanities Research Council of Canada and technical assistance was provided by the University of Hawaii. Gene Wilkening, Fritz Fliegel, John Walton, Morris Zelditch, and several anonymous reviewers commented on earlier drafts, and Bill Wood helped with computer analysis.

ernmentally provided services displace local dependence. (1971:11,7)

According to this view, local dependencies and inequalities are replaced by and, by inference, reduced by shifts to dependencies outside the village. Increasing dependency of the rural community on the larger society is required for development. In order to consider these different views I will examine village-based agrarian production systems within the Indian context.

UNEVEN DEVELOPMENT IN VILLAGE INDIA

Rural communities in underdeveloped nations are a basic feature of indigenous civilization. The Indian village continues to be a viable community for its inhabitants. A person's village still provides a source of self-identification, a nexus of activities, a stage for status, and an arena of conflict (Mandelbaum, 1970:421). The typical settlement pattern of most Indian villages is a nucleated dense cluster of houses surrounded by village fields. Villages may also include hamlets which may be composed of only one caste or one extended family (most often ex-untouchables). Other, much less common forms of settlement include linear patterns in which houses are strung along roads or canals, and homesteads or small cluster settlements found primarily among hill tribes. Even though villages remain important social, political, and economic units, they are obviously not isolated, self-sufficient, institutionally complete little republics (Mandelbaum, 1970; Desai, 1969).

Over 80% of India's population is rural and lives in more than half a million villages. Almost 70% of India's labor force is still employed in agriculture, a figure that has changed little over this century. The proportion of this agricultural labor force who were cultivators (mostly owners) and hired laborers had changed little until the 1960s. Since then, however, the proportion of male agricultural laborers as a percentage of all rural male workers increased from 15.8% in 1961 to 25.2% in 1971. The distinction between cultivators and laborers underestimates changes in the concentration of wealth. Still, inequality probably increased over that decade.[1]

Prior to the time of the 1970 survey, despite a rhetoric of equality, Indian governments' land and tenure reforms, taxation policies, and development expenditures were, at best, ineffectually applied to such ends (Lipton, 1968; Moore, 1966; Myrdal, 1968). Elites easily blunted redistributive programs to reap the benefits from rural development and agricultural extension efforts (Frankel, 1971; Parthasarathy and Prasad, 1974; Griffin, 1974; Gartrell, 1977). While some of the new green revolution technology (high-yielding seeds, chemical fertilizers, pesticides, etc.) is highly divisible, irrigation and mechanization usually are not (Gotsch, 1972). To the extent that government programs were concentrated in areas favorable to high production—particularly with assured water supply—regional and community differences in resource development levels were exacerbated. To the extent that the rich were disproportionately able to invest in the new technologies and disproportionately benefited from government programs, inequality in the distribution of resources and incomes within the community have also increased.[2]

Hypothesis 1a predicts that higher community development of resources will

[1] Indian census figures show that among all rural male workers, cultivators decreased from 61.1% in 1961 to 56% in 1971. Even sharper changes were observed for females. The proportion of female cultivators fell from 58.9% in 1961 to 32.7% in 1971 and agricultural laborers increased from 24.8% to 54.3%. Changes in inequality are only crudely reflected in these differences, since not all cultivators are owners and wealth and incomes vary within categories (particularly for cultivators). These changes might also result in part from the recruitment of labor by a more productive and expanded agriculture.

[2] The well-to-do are clearly in a better position to try the new green revolution technology (Gartrell, 1977). Indeed, yields of these new seeds were highly variable when compared to traditional varieties. In such circumstances poorer cultivators would quite reasonably "risk avert" in order to protect their subsistence, even when that meant foregoing long-term income gains (Weeks, 1970; Scott, 1976; Gartrell and Gartrell, 1979).

be associated with higher levels of inequality within the village (Dasgupta, 1975; Epstein, 1973; Myrdal, 1968). Given the conditions of capitalist production, the greater the surplus generated within the community, the greater the inequality in its distribution (Stanfield and Whiting, 1972; Moore, 1966; Maddison, 1971; Griffin, 1974; Mencher, 1974; Sharma, 1978).

Alternatively, the higher level of resources necessary to modernization is hypothesized to result in lower inequality within local economies (*Hypothesis 1b*). Several rationales may be offered in support of this prediction. Dependence and inequality may be "externalized" as rural communities are integrated into the wider society (Dalton, 1971). Modernization and community development can be viewed as forces which break the hold of traditional elites to bring an open, achievement-oriented stratification system (Singh, 1976). These kinds of changes can be seen to turn cumulative inequality into dispersed inequality (Beteille, 1969), and higher production may not result in greater inequality (Swenson, 1976).

If *Hypothesis 1a* is correct, and if government development programs and agricultural extension efforts merely reinforce the position of local elites, then we would expect that greater penetration of the village by the wider (capitalist) society would result in greater inequality within the village community (Sharma, 1978). This general prediction we label *Hypothesis 2a*. Alternatively, if some aspects of modernization and development reduce inequality (*Hypothesis 2b*), then we might also predict that, in some respects, villages which have closer connections to the wider society and higher government development inputs will exhibit lower inequality (Rao, 1966; Beals, 1974). Assuming an isomorphism of effects at the community level and those at the nation-state level, welfare spending on "human capital development" may reduce inequality (Snipp and Summers, 1980; Adelman and Morris, 1973). There are, however, aspects of agrarian communities that have little parallel at the nation-state level. Precapitalist aspects of local village economies such as caste differences and the village exchange system which they support may be associated with lower inequality.

PRODUCTION AND EXCHANGE

Village agricultural production and distribution systems appear designed to insure subsistence at the price of high material inequality (Scott, 1976; Sharma, 1978). Local *jajmani* exchange involves inherited linkages between food-producing families and those who supply them with labor, artisan manufactures, and other services (Mandelbaum, 1970). Relationships were traditionally governed by caste *panchayats* (councils), and involved ceremonial obligations which reinforced ritual differences in the caste system. Exchange rates were usually regulated by custom, but actual payments might vary with supply and demand in terms of need, abundance, generosity, and the quality of the work performed. Jajmani ties involved general obligations of mutual support, and political conflict often involved the mobilization of jajmani associates. Jajmani exchange therefore united villages by organizing castes vertically across class lines.

Such systems are obviously exploitive. Surpluses are monopolized by the landowning elite. Yet one reason for the tenacity of the system may lie in the subsistence "guarantees" it offered to landless, subordinate laboring castes (Epstein, 1973). Since the rich were obliged to provide for the subsistence of the poor, such a system may actually be less exploitive than a more modern, market-oriented, capitalist system (Scott, 1976:38). The latter generally involves exchange in cash rather than in kind, and this increases inequality where prices rise faster than wages (Myrdal, 1968; Griffin, 1974). Monetization may also depersonalize jajmani relationships, restrict side-payments, and generally reduce employment security (Frankel, 1971; Freeman, 1977; Sharma, 1978). *Hypothesis 3a* predicts that the maintenance of a traditional division of labor and traditional patterns of exchange will lower inequality. To the extent that precapitalist elements of local social and economic structure are strengthened and maintained, this will

limit the effects of penetration by the wider society (Amin, 1976).

Alternatively, *Hypothesis 3b* predicts that traditional social structures act as an impediment to development and to the reduction of inequality. This is the tack that most development theorists have taken (for a review, see Bloomquist, 1980). Land-owning local elites clearly have good economic reasons for abandoning traditional exchange systems. The conditions of capitalist wage-labor clearly offer possibilities for greater economic exploitation where there is a surplus of labor and it is poorly (or not at all) organized. There are reasons to give up the assured labor supply tied to land-owners and the easier social control and political support that the jajmani system may bring. The State may provide control functions (police). Yet the prestige of the elite and the social order of the village are both considerably enhanced by the ritual order which reinforces the jajmani system. Based on the proposition that status inequality ritualized in the caste system reinforces material inequalities, *Hypothesis 4* predicts that these two dimensions of inequality are positively related. While this hypothesis is of somewhat less substantive interest (it does not differentiate the views outlined above), it is necessary to control for caste ritual inequality when examining the influence of other factors.

To summarize, these hypotheses have been organized into two, oversimplified, general perspectives and their differences have been maximized in order to facilitate exploratory analysis. The dependency model—hypotheses 1a, 2a, and 3a—sees government inputs as directed towards increasing the development of capitalist agriculture. Higher levels of development (surplus) and higher penetration (inputs) by the wider society are positively related to inequality. Also, precapitalist elements of local economic and social systems may act to limit the penetration of the wider society that decreases the autonomy of the local community. Stronger traditional jajmani relationships would therefore be associated with lower inequality. Against this, the modernization model—hypotheses 1b, 2b, and 3b—predicts opposite effects for the same antecedents. Development and modernization result in lower inequality. Higher government inputs and the increasing integration of the village into the wider society are necessary for higher levels of development. By increasing production, breaking down traditional social structures (seen as ascriptive), and replacing them with new achievement-oriented opportunities, modernization lowers inequality within these agrarian communities.

SAMPLING

The south-central Indian state of Andhra Pradesh had a population of about 43.5 million in 1971. The state includes both relatively well-developed, irrigated rice cultivation in the coastal Andhra region and extensive areas of dry cultivation in the Telangana region centered on the inland Deccan plateau. It thus provided the variability in village development levels necessary to an examination of the effects of modernization and dependence at the community level. In designing the sampling frame for the 1970 survey, the 20 districts within the state were ranked on a six-year average (1961–1967) of gross farm income per cultivating household. The first, fourth and ninth ranked districts were selected from within each of the state's two principal regions (Andhra-Rayalseema and Telangana). This procedure for the sampling of districts was designed to insure regional representation and to include districts at varying levels of agricultural productive capacity. The Taluks which made up these six districts (West Godavari, Guntur, and Chittoor; Nizamabad, Karimnagar, and Mahbubnagar) were then divided into high and low irrigation strata based on a three-year average (1964–1967) of gross irrigated area divided by gross cultivated area.[3] In each of the six districts, 14 villages with popu-

[3] More detailed information was available in West Godavari where the irrigation lists provided a suitable sampling frame which stratified villages directly. Stratification into irrigated and dry-cultivation areas was effective in sampling from districts like West Godavari and Nizamabad. In Districts that did not have both wet and dry areas such as Mahbubnagar (dry) and Guntur (irrigated), this stratification mattered relatively little.

lations between 200 and 5,000 were selected in proportionate random samples from the high and low irrigation areas. This yielded a total sample of 84 villages. The sampling design over-represented relatively well-developed villages, since a constant number of villages was drawn from all districts and richer districts had fewer villages. As was the case with stratification of districts and Taluks on the basis of development level (production levels and irrigation), this strategy was adopted to maximize regional representation and to expand the variance in developmental conditions within the sample.

Village-level information was obtained in two ways. First, community characteristics were sought directly in interview with knowledgeable informants based upon their records.[4] Second, properties of the village were estimated from interviews with a sample of household heads. In each of the 84 villages, households were selected in proportionate random quota samples from three crude occupational strata: cultivators, agricultural laborers, and artisans and others. Those employed outside of agriculture were included in the latter sampling strata. As villages increased in size, sample size was increased while the sampling fraction was decreased. Samples ranged from 15 out of 40 households in the smallest village to 60 out of 863 households in the largest. This procedure yielded a total of 2,040 interviews with heads of households.

Pretesting was conducted during February and March of 1970 in three villages (two in Mahbubnagar District, one in Krishna District). These three villages were not included in the final sample of 84 villages (Gartrell, 1973). Data collection began in West Godavari in late March and finished in Mahbubnagar in September. Field work was timed to minimize problems that might have been created by labor migration (particularly in Andhra), a phenomenon that appears to have increased over the past decade.

MEASUREMENT OF COMMUNITY DEVELOPMENT AND DEPENDENCE

Community development levels, the integration of the village into the wider society, the penetration of the village through government inputs, and any consequent dependence or lower autonomy on the part of the community, have been discussed as analytically independent constructs. They are obviously closely related. Attempts to operationalize them revealed a number of points at which they generally overlapped. For example, government material or personnel inputs constituted penetration of the village, but also reflected the village's integration into the wider society. In a sense, the village was dependent on government agencies for continued inputs. The present level of village resources, for example literacy or electrification, was a fairly direct consequence of past government spending. Such multiple meanings necessitated considerable caution in discussing operational definitions and in interpreting results of multivariate analysis.

These connections suggest that development and dependence will be positively related. At the community level, dependence does not appear to have inhibited the accumulation of resources. Larger villages with higher levels of both material and human resources, more complex occupational structures, and more commercialized agriculture would also be expected to have closer linkages to the wider society. They would probably receive more development inputs and would exhibit internal exchange systems that reflect greater penetration. These predictions do not distinguish between modernization and dependence theories of center-periphery relations. They do, however, give a working guide to the selection and interpretation of indicators.

[4] Four long interviews with knowledgeable informants were conducted in each of the 84 villages. These included the Patwari, the land and tax record keeper; the local community development agent (Village Level Worker); a Panchayat (village council) official, usually the Sarpanch (head of the council) or the council clerk; and the Taluk Statistical Clerk. Interviews were based on their written records. Research design benefited from the experience of the Diffusion of Innovations Project directed by Everett Rogers (Roy et al., 1968; Fliegel et al., 1968). He generously made the data available for analysis prior to the present survey. Frederick Fliegel, C. Lakshmanna, and Lalet Sen provided helpful comments and data collection was facilitated by Osmania University and the Government of Andhra Pradesh.

In operationalizing "development," the level of productive resources (material) was measured as the average agricultural capital (land, tools, livestock, buildings) per household estimated from sample survey results. Average gross crop income per household was aggregated from survey questions as to crops sown, acres planted, yields obtained, amount sold, and prices received. Unsold produce was valued at harvest prices. Average household incomes were aggregated from income estimates obtained for each employed household member. The level of technological development in the village was indexed by the percentage of households electrified (village records) and the complexity of village occupational structure was indicated by the proportion employed in agriculture (survey estimates). An index of housing quality, estimated from sample survey results, operationalized an important aspect of the level of living; and the percentage of literate household heads was used as a measure of "human capital." The index of housing quality and the measure of literacy were adapted from Roy et al. (1968). The housing index summed the presence of modern materials and construction design (windows with shutters, floors of something besides dirt, something other than a thatched roof, etc.). Literacy was judged by responses to questions asking if the head of the household could write a letter and read a newspaper. Population size was introduced primarily as a control, but it also crudely reflects human resources available in the village. Larger villages may also be relatively more institutionally complete and so less dependent on regional markets, other villages, or the larger society in general.

As is typical of Andra Pradesh, villages were fairly large (averaging 1,436; s=1,136), relatively poor (farm capital averaged Rs. 21,000, about $2,800; s = 13,000) and had simple occupational structures (88.6% employed in agriculture, s = 10.7). Household incomes were low (averaging Rs. 2,152) and few households had electricity ($\bar{x}$ = 11.2%, s = 14.6). An average of only 30% of household heads were literate (s = 16.5). As expected, these different aspects of development were positively intercorrelated (negative with percent in agriculture). For example, higher levels of capital accumulation (surplus) were associated with higher literacy (r = .516), electrification (r = .503) and housing quality (r = .394). Using a one-tailed test with a sample size of 84 communities, correlations greater than or equal to 0.181 are significant at the .05 level (0.253 at the .01 level). Since electrification and literacy reflect inputs by the state government, such results reflect past concentration in development efforts. (See Appendix for full zero-order correlation matrix.)

Further evidence of concentration of development efforts can be seen in two additional measures: the rupee total of government agricultural inputs per year (seeds, fertilizers, pesticides, and the credit to purchase such items) tabulated from Village Level Worker records, and change agent contact per year per village household (survey estimates).[5] Villages had received an average of Rs. 16,500 per year (s = 25,100), with inputs concentrated in some villages and distribution highly concentrated within villages as well (Gartrell, 1976). Households averaged about 20 contacts per year with extension personnel, primarily the Village Level Workers stationed in the villages. This too varied widely across villages (s = 18.1) and was concentrated within them (Gartrell and Gartrell, 1979). The observation that government developmental inputs have been concentrated in wealthier villages was reflected in the pattern of intercorrelations. For example, agent contact

[5] Agents included primarily the Block Development Officer, Agricultural Extension Officers, and the Village Level Workers. State community development efforts in Andhra Pradesh are administered through Community Development Blocks made up of about 50 villages. The organization of Village Panchayat, Samiti Panchayat (Block level), and Zilla Parashad (District level) paralleled but did not completely coincide with the separate Federal revenue and justice administration. Federal administration divides the State into Districts divided into Taluks or Tahsils divided into Revenue Circles made up of Revenue Villages. Since the latter generally better approximated the separate residential collectives regarded by their inhabitants as a "village," it was used as the sampling unit rather than the Panchayat Village, which often included more than one village if they were small.

was positively correlated with levels of village agricultural capital ($r = .434$) and literacy ($r = .343$), as was the level of agricultural inputs ($r = .306$ and 0.387).

A communications resource index (survey estimates of newspapers, radios, and vehicles per 100 households) represented a different aspect of center-periphery relations. As a related negative measure of the penetration, an index of village isolation was constructed as the sum of the distance in miles to the nearest railway station, all-weather road, and bus stop. Both of these indicators described the villages as only weakly linked to the wider society. The communications index ($\bar{x} = 16.3$, $s = 16.1$) may have underestimated these linkages, since radios, newspapers, and vehicles can all be shared by a number of users. Similarly, the isolation index ($\bar{x} = 23.9$, $s = 22.0$) may have overestimated the difficulty of urban contact, since access to any form of transportation is probably sufficient. The two measures were negatively correlated ($r = -.444$) with better communication resources found in richer villages ($r = .372$) that had more complex occupational structures ($r = -.313$). Higher communications resources were also observed in villages with higher literacy ($r = .540$) and more widespread electrification ($r = .572$), as well as those with high levels of agricultural and change-agent inputs on the part of the government ($r = .321$ and $.287$). The pattern of zero-order effects was weaker but similar (with the opposite sign) for the isolation index.

Dependence upon external markets was indexed by the percentage of crops sold (by value) in the crop year preceeding the survey. Greater production for sale rather than for consumption or barter may have increased the dependence of village producers on markets over which they had little control. The monetization of exchange of goods and services within the local jajmani system also reflects penetration by the wider society. It was measured as the percentage of jajmani relationships in surveyed households where the majority of exchange was reported to be in kind. This indexes the presence of precapitalist exchange. Villages rarely contained all the relevant jajmani services. The degree of local intervillage autonomy was indexed as the percentage of all jajmani relationships where both parties lived within the village.

The penetration of these villages by the metropolitan cash economy was only moderate and varied widely. Only 47.6% of crops were sold on the average ($s = 16.4$). This index of commercialization in agriculture was correlated with the level of farm capital ($r = .439$), crop income ($r = .308$), average household incomes ($r = .463$), communications resources ($r = .387$) and electrification ($r = .542$). However, the percentage of crops sold was linearly unrelated to jajmani exchange in kind, the percentage employed in agriculture, or the percentage of jajmani services within the village. The degree of monetization of local exchange was also modest. An average of 51.7% of jajmani relationships were primarily in kind ($s = 19.2$). Exchange in kind was somewhat more likely in more occupationally homogeneous ($r = .248$) or relatively isolated villages ($r = .289$) and those with lower literacy ($r = -.248$) and poorer communications ($r = -.211$). However, exchange in kind was not significantly correlated with the level of wealth or income in the village. Jajmani exchange took place primarily within the village ($\bar{x} = 79.4\%$, $s = 27.4$). Villages with more exchange within the village also had a higher percentage of exchange in kind ($r = .429$). Larger villages ($r = .245$) and those with better housing ($r = .319$) were less dependent on procuring services from beyond their borders. Otherwise, jajmani autonomy with respect to nearby villages appeared to be unrelated to other aspects of village autonomy or levels of development.

To summarize, larger, wealthier villages with more complex occupational structures tended to have better housing, more electrification, and higher literacy. Not surprisingly, these "better developed" villages were less isolated, had better communication links with the larger society, and had received more development inputs from the government. They had a more commercially oriented agriculture, and somewhat more monetized exchange within the village. These results replicate

Adelman and Dalton's (1971) factor analysis of 108 villages and Dasgupta's (1975) results from a sample of 126 communities. Government inputs have been concentrated in more developed villages and community development and dependence generally tend to be positively associated.

THE MEASUREMENT OF INEQUALITY

Cross-national studies of inequality have found it difficult to obtain accurate measures (Kuznets, 1963; Kravis, 1962; Rubinson, 1976; Paukert, 1973). While the use of household income distributions represented a substantial improvement over intersectoral estimates, incomes are very difficult to measure in only partially commercialized, underdeveloped, agrarian societies. Unfortunately, errors in measuring income inequality for nations are probably negatively related to development (and perhaps positively to dependence). This may in part account for the relatively low proportion of variance in income inequality accounted for by either of these factors.

As indicated above, three different aspects of wealth and income were measured in the 84 sample surveys: agricultural capital, gross crop income, and household income. In an attempt to improve upon the common practice of studying rural inequality from distributions of land ownership (or worse, cultivation) in acres, the interviews obtained self-reports from household heads as to the value of their land (by plot and type of irrigation). To this were added inventories of the value of machinery and tools, livestock, and farm buildings, to form an index of farm capital (Roy et al., 1968). The Gini coefficient of concentration for the distribution of agricultural capital across households was used as the index of inequality.[6]

The measurement of income presented even more problems than the gentle obfuscation that must have met questions about assets. Most of the labor force did not work for cash wages or salaries. Farm records were rare, and costs of production were difficult to estimate (Epstein, 1973). The wages of agricultural laborers and others may sometimes represent net incomes. At other times nonmonetary wages (meals, food, clothing, loans, gifts) typical of traditional jajmani relationships make even laborers' income difficult to quantify. A second strategy therefore was to ignore labor incomes, variable production costs for cultivators, and the production of goods other than crops and to focus on gross crop incomes. The Gini coefficient for the distribution of gross crop income across households was used to index inequality.

Estimates of total household incomes by household heads suffered all the problems of attempting to combine cultivators' incomes with those of laborers. Neither kept accounts, and both had a tendency to remember cash and to forget payments or income in kind. These estimates should be viewed as best guesses, and they may in some instances be deliberately fabricated. The use of all three measures of inequality has one additional advantage. Their interrelationship is relatively straightforward. If laborers' incomes vary relatively little, household income is largely a function of income from cultivation ($r = .606$). That in turn is dependent on inputs of agricultural capital ($r = .445$).

Another potentially important dimension of inequality in Indian agrarian communities involves religious status differences in ritual purity across hierarchically ranked castes (*jatis*). Ritual inequality was of particular interest because it reinforces jajmani exchange as well as legitimizing material inequalities. While the measurement of ritual purity in different villages with no direct interaction was hazardous at best (Roy et al., 1968; Morrison et al., 1976; Marriott, 1965) a scale was developed based on Lakshmanna's (1973) field work in 11 districts

---

[6] The Gini coefficient may be relatively insensitive to differences at the extremes of the distribution. At the risk of complicating the analysis, quintile or quartile measures might have been added to allow examination of the pattern of concentration or distribution for different factors. However, given the small size of village samples (15 to 60 households), even quartile estimates would have been based on very few observations.

of Andhra Pradesh (Gartrell, 1973). Based upon ritual status assigned to jatis after field work was completed, jati ritual inequality was measured as the Gini coefficient for the distribution of status across households sampled within each village.

The 84 communities varied widely on all four measures of inequality (see Table 1). By way of comparison, the average inequality in the distribution of income for the 47 nations for which Rubinson (1976) could find suitable data was slightly lower and the standard deviation was actually smaller (9.5 versus 11.4). Relatively high variation across communities within a single nation was also observed by Stanfield and Whiting (1972) for Brazil. Inequality may vary considerably within underdeveloped countries.

Table 1. Inequality in the Distribution of Farm Capital, Income, and Jati Ritual Status

Inequality Measures	Zero-order Correlations 1	2	3	Mean	St. Dev.
1. Jati ritual status	—			25.72	11.91
2. Farm capital	.342	—		70.25	13.26
3. Gross crop income	.301	.795	—	69.79	12.35
4. Household income	.077	.198	.308	46.15	11.35

NOTE: Gini coefficients were multiplied by 100 to represent the percentage of area between the Lorenz curve and the line of equality.

Inequality in the distribution of both gross crop income and agricultural capital were much higher on the average than household income inequality. Yet the latter estimate was very similar to that reported in a number of other studies (for comparisons see Swensen, 1976). The average level of jati ritual status inequality was relatively low because the highest ranking Varnas were numerically very small in Andhra Pradesh. Sudra and exuntouchable castes formed the bulk of most villages. Perhaps for that reason, ritual status differences may have been less important than has often been assumed (Sharda, 1977). Ritual status inequality had only modest positive correlations with other dimensions of inequality (Table 1) and was virtually orthogonal to household income inequality. As expected, crop income inequality was highly dependent upon capital inequality (r = .795). Household income inequality was moderately related to crop income inequality, but the effects of farm capital inequality on household income inequality appeared to be largely indirect through gross crop income inequality.

MULTIPLE REGRESSION RESULTS

Three multiple regression equations are reported below, one for each inequality measure. All 14 measures of development and dependence, as well as the relevant inequality measures, were included as potential predictors. For each different Gini coefficient the average level of the distribution was included. Thus average farm capital was used as a potential predictor of farm capital inequality, average crop income was used for crop income inequality, and average household income was included for household income inequality. In order to simplify results, reduced-form equations are reported.[7]

Seven of the 14 variables had significant independent effects on farm capital inequality (Table 2). As predicted in Hypothesis 4, jati ritual status inequality had significant positive effects on equality in the distribution of agricultural capital (b = .365). Indeed, control for other community characteristics had little influence upon its observed effects (beta = .328; r = .342). The higher the proportion of jajmani exchange in kind, the lower the farm inequality in the village (b = −.164). The maintenance of traditional patterns of exchange had strong negative zero-order effects on farm capital inequality (r = −.424) that were reduced by the addition of controls (beta = −.236). These results supported Hypothesis 3a. So too did the

[7] Each bivariate relationship between the predictors and the three inequality measures was tested for nonlinear components. In no case did simple linearizing transformations improve prediction. This was also the case for the effects of one dimension of inequality upon another. An exploratory analysis was also initiated to test the covariance (interactive) hypothesis that, given different conditions within the village, government inputs would have different effects on inequality. Little systematic evidence was found to support this hypothesis, perhaps because of the confounding of resources and inputs.

Table 2. Agricultural Capital Inequality: Multiple Regression Results

Variable	b	Std. Error	Beta	r
% Exchange in kind	−.164	.065	−.236	−.424
Ritual inequality	.365	.097	.328	.342
% Employed in agriculture	−.357	.118	−.287	−.328
Housing quality	−3.437	1.190	−.270	−.204
Communications (/100 H.H.)	.367	.088	.284	.247
% Literate	−.208	.087	−.260	.088
Gov't agricult. inputs (Rs. 1000)	.107	.050	.202	.219
$R^2 = .453$	$R_c^2 = .403$	$S_e = 10.25$	$F = 8.99$	

NOTE: All partial slopes are significant at the .05 level (df = 1, 76). Beta refers to the standardized partial regression coefficients.

negative effects of a more homogeneous occupational structure (b = −.357). To the extent that a more heterogeneous occupational structure reflects higher levels of community development, these results also lend support to Hypothesis 1a. The greater the occupational complexity, the greater the inequality.

Similar effects were observed for the communications index (b = .367). The higher the level of development resources and the greater the village's exposure to urban (national) society, the greater the concentration of farm capital. These results supported Hypotheses 1a and 2a. Higher governmental agricultural inputs were also associated with greater concentration of agricultural capital (b = .107). In general, material inputs, like the accumulated level of material resources, appeared to be associated with concentration in the distribution of farm capital.

The opposite effects were observed for both housing quality and literacy. The index of housing quality had distributive effects on farm capital inequality (b = −3.437), as did the level of literacy in the village (b = −.208). When development was seen in terms of human welfare (human capital), a higher level of development was associated with lower inequality. The small positive zero-order effects of literacy (r = .088) appeared to be suppressed by communication resource levels (these two predictors had a high positive correlation, r = .540). In the case of literacy, human capital inputs appeared to have distributive effects only when the concentrating effects of communications were controlled. The observed effects of housing quality on farm capital inequality (r = −.204) were also slightly suppressed (beta = −.270). When the step-wise regression procedure was set to include only variables significant at the .05 level, none of the other seven variables had significant partial effects. Neither did they have significant zero-order effects.[8]

As expected, inequality in the distribution of gross crop income was largely a function of inequality in the distribution of agricultural capital. Its zero-order effects (b = .740) were reduced only slightly by the addition of controls (Table 3). Those factors which directly affected farm capital inequality thus had similar, smaller indirect effects upon crop income inequality through farm capital inequality.

Table 3. Inequality in the Distribution of Gross Crop Income: Multiple Regression Results

Variable	b	Std. Error	Beta	r
Farm capital inequality	.689	.061	.739	.795
Isolation	−.091	.034	−.162	−.214
Agent contact	−.169	.045	−.247	−.267
Crop income (Rs. 1000/H.H.)	6.842	2.760	.164	−.122
% Employed in agriculture	−.171	.075	−.148	−.387
$R^2 = .725$	$R_c^2 = .707$	$S_e = 6.683$	$F = 41.14$	

NOTE: All partial slopes are significant at the .05 level (df = 1, 79). Beta refers to the standardized partial regression coefficients.

[8] When inclusion levels for the reduced-form equation were lowered to F=3.0, and average farm capital per household was forced into the equation, several predictors could be added. The pattern of results remained unchanged. Material inputs or resources (agricultural capital, electrification) were associated with increased concentration. Higher personnel inputs or human resources (agent contact, population size) had distributive effects. These additional predictors had nonsignificant zero-order correlations with inequality; but a complex pattern of suppressor effects (material factors suppressing human and vice versa) yielded partial effects significant at the .10 level. A higher percentage of jajmani exchange within the village was associated with greater inequality, perhaps because of the greater control it offered. These five predictors added significantly (F = 3.99; df = 5.71) to R^2 (.573). Otherwise, they had little effect upon the results reported in Table 3.

Net of the effects of farm capital inequality, gross crop income inequality was also a negative function of four additional variables. These results followed a pattern similar to that observed for inequality in the distribution of farm capital. The more isolated the village and presumably the less penetrated it was by the wider society, the lower the inequality in gross crop income (Hypothesis 2a). A more homogeneous occupational structure was associated with lower inequality (Hypothesis 3a), but greater contact with government change agents was associated with lower crop income inequality. This corresponds to the community modernization prediction (Hypothesis 2b) and provides further evidence that human resource inputs had redistributive effects.

The level of community development, as indicated by average crop income per household, had significant positive partial effects on crop income inequality. These results support Hypothesis 1a, although the observed zero-order correlation was negative. The first-order partial correlation with farm capital inequality controlled was .133. Again, the level of material resources had positive effects on inequality.

The average level of household income in the village was the most important determinant of household income inequality. It accounted for over one half of the variance explained (Table 4). The effects of the average levels of inequality variables were thus consistently positive, even though these effects were suppressed for both farm capital and crop income. The results for household income provided the strongest support for Hypothesis 1a, linking the level of development with greater inequality. The somewhat more modest effects of inequality in the distribution of gross crop income (b = .422) may reflect measurement problems in estimating household incomes (and also crop incomes).

The level of household income and crop income inequality both acted to suppress the negative effects of communication resources upon household income inequality (b = −.241). The negative first-order partial effects (r = −.136) with agricultural capital controlled showed a sign change from the observed positive zero-order effects (r = .103). Second-order partial effects (r = −.251) with the additional control of crop income inequality showed that it too suppressed the negative effects of communication resources. These effects were quite the opposite of those observed for communication resources in the equation predicting inequality in the distribution of agricultural capital (Table 2). As reflected in the communications index, the closer integration of the village into the wider society was associated with high concentration of farm capital and lower inequality in the distribution of household income. However, these latter distributive effects were observed only when the income-concentrating effects of the level of household income and inequality in the distributive of crop income were controlled.

Table 4. Inequality in the Distribution of Household Income: Multiple Regression Results

Variable	b	Std. Error	Beta	r
Household income (Rs. 1000/H.H.)	4.958	.890	.507	.541
Crop income inequality	.422	.085	.459	.308
Agent contact	.223	.060	.356	.303
Communication (/100 H.H.)	−.241	.068	−.342	.103
$R^2 = .488$	$R_a^2 = .462$	$S_e = 8.326$	F = 18.81	

NOTE: Inclusion levels were set to the .05 level of significance, but all partial slopes are significant at the .01 level (df = 1, 79). Beta refers to the standardized partial regression coefficients.

Agent contact had positive effects on household income inequality (b = .223). These results were exactly the opposite of those which agent contact had upon inequality in the distribution of gross crop income (Table 3). In this case the observed effects (r = .303) were consistent with the partial effects (beta = .356). They tend to support Hypothesis 2a and run counter to the observation of distributive effects for human inputs or resources. To the extent that change agent contact has been directed almost exclusively at raising agricultural productivity, it has concentrated only on cultivators' incomes. While laborers' wages may also benefit from production gains (Swensen, 1976), other

income streams were less likely to be affected.

DISCUSSION

As the dependency model of center-periphery relations predicted, material inputs and resources had concentrating effects on the distribution systems within agrarian villages (Hypotheses 1a and 2a). Contrary to much of what is written about the Indian caste system, the maintenance of a traditional division of labor and jajmani exchange generally appeared to result in lower inequality (Hypothesis 3a). The single significant exception to these predictions among the observed zero-order correlations was the negative association between the index of average housing quality in the village and the concentration of farm capital. The higher the level of human welfare in the village, the lower the inequality. These results supported the community modernization predictions (Hypothesis 1b). The community development model received additional support from the negative partial effects of literacy on the concentration of farm capital and the negative partial effects of communication resources on household income inequality. In both cases, the observed zero-order effects, although nonsignificant, were positive. Only when the hypothesized effects of the dependency model were controlled did these distributive effects predicted by the modernization model become apparent.

A closer look at between-village differences in literacy rates suggests that they may have been partially a function of the timing of inputs of primary education. Communities which received educational facilities earlier would have had a high proportion of literates among household heads. Education itself reflected this pattern. Average inequality in the distribution of years completed by adult males was high ($\bar{x} = 73.4$) and varied widely ($s = 16.0$). The distributive effects of inputs of primary education were evident in the large negative correlation between average levels of education and inequality in its distribution ($r = .625$).[9]

Not all government inputs resulted in greater inequality, although material resource inputs and higher levels of income did increase inequality. The problem for future development may be that rural educational opportunities have not been extended beyond elementary schooling. When all adults have had the advantages of existing opportunities, and differences in timing of inputs are no longer relevant, the distribution of education may become more closely linked to the distribution of wealth and income. Unfortunately, the distributive effects of past efforts may be transitory.

Rural labor migration may also influence these and future results. The timing of field work was designed to minimize the absence of laborers, but future data collection efforts might better address the problem more directly. The rate of temporary rural labor migration appears to be increasing, and if migrant labor were omitted from samples of rural communities, the poor would be underrepresented. Village resource levels would be overestimated and inequality would be underestimated. A negative bias might thus be introduced into the association between development levels and inequality. For the present results, such a bias would only have rendered the observed results a more conservative estimate. However, the development of different forms of labor requires further research.

There are also problems in interpreting the observed negative association between housing quality and inequality. Levels of living in rural areas may have declined prior to the 1970 survey (Indian Institute of Public Opinion, 1968:45; Dandekar and Rath, 1971). If we assume such a trend, then decreasing welfare levels could be linked to greater inequality. A negative relationship between levels of living and inequality might thus be described by an impoverishment hypothesis.

[9] A similar pattern of relationships was observed for knowledge of new agricultural practices. The level of information in the village was strongly negatively correlated with inequality in its distribution ($r = -.779$). Information was widely diffused, probably through a two-step flow of communication, and this too indicates a degree of success for government agricultural extension efforts.

It would generate predictions no different from those of the community modernization model (Hypothesis 1b). Only data on temporal change for individuals and communities could bring evidence to bear on these different interpretations.

The causal ordering of effects is much more complicated than that represented in the regression analysis presented above. The integration of the village within the wider society (dependence, penetration) describes the relationship between communities and larger social units. In this sense, center-periphery relationships are exogenous to the community itself and therefore antecedent to inequality within the village. The accumulation of surplus or resources and the division of labor can also be seen as determinants of the distribution of resources. The analysis above assumes this directionality of effects. Yet it can be plausibly argued that inequality influences subsequent community development as well as the integration of the community into the wider society. Longitudinal data would be necessary to investigate such feedback effects.

Finally, from the modernization perspective it can also be argued that the association of higher levels of development with greater inequality is itself a temporary phenomenon. When only part of the community is exposed to modernization, inequality is increased. When all sectors of the village enter the modern cash economy, modernization differences between villages are reduced and the positive effect of modernization on inequality might disappear.

However, local village economies in Andhra Pradesh, even in the relatively backward Telangana region, have been partially monetized for a long time (Iyengar, 1931). Change appears to be slow. Laborers resist the transition to cash because it erodes their wages and working conditions (Epstein, 1973; Frankel, 1971; Freeman, 1977; Sharma, 1978). In the sample of 84 villages, the level of agricultural capital accumulation, average crop income, and average household income were virtually orthogonal to the degree of monetization of jajmani exchange (correlations of .024, .091 and .001, respectively). All three of these indicators of development levels were significantly positively correlated with the percentage of crops sold, but this commercialization of agriculture had no significant independent effects on inequality when other factors were controlled. The observed positive effects of development on inequality do not appear to be temporary. Yet until more direct evidence of change can be marshalled, any conclusion must be cautious.

APPENDIX Zero-Order Correlations

Variables	10	11	12	13	14	15	16	17	18	19
1. Income inequality	—	.248	−.313	.201	−.218	−.143	−.130	−.008	−.135	−.248
2. Crop income inequality	.308	—	−.211	.289	−.167	.429	.134	−.088	−.107	−.171
3. Farm capital inequality	.198	.795	—	.444	.572	.065	.237	.387	.310	.321
4. Jati ritual inequality	.077	.301	.342	—	−.364	.130	−.084	−.130	−.222	−.297
5. Income/H.H.	.541	.106	−.055	−.143	—	.030	.286	.542	.364	.290
6. Crop income/H.H.	.448	−.122	−.252	−.205	.609	—	−.009	−.042	.244	−.065
7. Farm capital/H.H.	.534	.045	.017	−.110	.606	.445	—	.095	−.094	.156
8. % Literate	.336	.194	.088	.037	.511	.184	.516	—	.231	.224
9. Housing quality	.103	−.155	−.104	−.010	.306	.204	.394	.225	—	.559
1. Income inequality	−.163	−.095	.103	−.135	.214	.007	.303	.224	.013	.162
2. Crop income inequality	−.387	−.482	.322	−.214	.157	−.075	−.267	.087	.112	.232
3. Farm capital inequality	−.328	−.424	.247	−.057	.083	−.026	−.158	.062	.010	.219
4. Jati ritual inequality	.041	−.160	−.038	.047	−.144	.086	−.126	−.146	.041	.092
5. Income/H.H.	−.226	−.001	.386	−.255	.484	.062	.331	.463	.164	.189
6. Crop income/H.H.	.021	.091	.129	−.005	.183	−.142	.393	.308	−.122	−.058
7. Farm capital/H.H.	−.098	.024	.372	−.185	.503	.098	.434	.439	.360	.306
8. % Literate	−.367	−.248	.540	−.372	.563	.044	.343	.272	.345	.387
9. Housing quality	−.180	.123	.330	−.141	.487	.319	.166	.296	.380	.206

ADDITIONAL VARIABLES:
10. % Employed in agriculture 12. Communications 14. % Electrifications 16. Agent contact 18. Population

APPENDIX—Continued

11. % Exchange in kind	13. Isolation	15. % Exchange within the village	17. % Crops sold	19. Government agricultural inputs

NOTE: Correlations among variables 10–19 appear above the diagonal in the top right section of the matrix. Correlations among variables 1–9 are below and to the left. The 9x10 matrix that forms the bottom half of the table contains the correlations between these two sets of variables (1–9 and 10–19).

REFERENCES

Adelman, Irma and George Dalton
1971 "A factor analysis of modernization in village India." Pp. 492–517 in George Dalton (ed.), Economic Development and Social Change. Garden City, N.Y.: Natural History Press.

Adelman, Irma and Cynthia Morris
1973 Economic Growth and Social Equity in Developing Countries. Stanford: Stanford University Press.

Amin, Samir
1976 Unequal Development. New York: Monthly Review Press.

Beals, Alan R.
1974 Village Life in South India. Chicago: Aldine.

Beteille, Andre
1969 Caste, Class and Power. Berkeley: University of California Press.

Bloomquist, Leonard E.
1980 "Persistence and change in India's village economies: a review of Indian community studies since 1960." Midwest Sociological Society Meetings, Milwaukee, Wisconsin.

Bornschier, Volker, Christopher Chase-Dunn, and Richard Rubinson
1978 "Cross-national evidence of the effects of foreign investment and aid on economic growth and inequality: a survey of findings and a reanalysis." American Journal of Sociology 84:651–83.

Chase-Dunn, Christopher
1975 "The effects of international economic dependence on development and inequality: a cross-national study." American Sociological Review 40:720–38.

Chilcotte, Ronald and Joel C. Edelstein (eds.)
1974 Latin America: the Struggle with Dependency and Beyond. Cambridge, Mass.: Schenkman.

Dalton, George
1971 Modernizing Village Economies. Reading, Mass.: Addison-Wesley.

Dandekar, V. M. and Nilakantha Rath
1971 Poverty in India. Poona: Indian School of Political Economy.

Dasgupta, Biplab
1975 "A typology of village socio-economic systems." Economic and Political Weekly 10:1395–414.

Davis, Arthur K.
1971 "Canadian society and history as hinterland versus metropolis." Pp. 6–32 in Richard J. Ossenberg (ed.), Canadian Society: Pluralism, Change and Conflict. Scarborough: Prentice-Hall.

Desai, A. R. (ed.)
1969 Rural Sociology in India. Bombay: Popular Prakashan.

Dos Santos, Theotonio
1970 "The structure of dependence." American Economic Review 60:231–6.

Epstein, T. Scarlett
1973 South India: Yesterday, Today and Tomorrow. New York: Holmes and Meier.

Fliegel, Frederick C., Prodipto Roy, Lalit K. Sen, and J. E. Kivlin
1968 Agricultural Innovations in Village India. Hyderabad: National Institute of Community Development.

Frank, Andre Gundar
1966 "The development of underdevelopment." Monthly Review 18:17–31.
1967 "Sociology of development and underdevelopment of sociology." Catalyst 3:1–67.

Frankel, Francine R.
1971 India's Green Revolution: Economic Gains and Political Costs. Princeton: Princeton University Press.

Freeman, James M.
1977 Scarcity and Opportunity in an Indian Village. Don Mills, Ontario: Cummings.

Gartrell, John W.
1973 "Development and social stratification: the structure of inequality in Indian peasant communities." Ph.D. dissertation, Dept. of Sociology, University of Wisconsin.
1976 "Inputs and inequality: towards an assessment of community development in Andhra Pradesh villages." Kerala Sociologist 4:152–66.
1977 "Status, inequality and innovation: the green revolution in Andhra Pradesh, India." American Sociological Review 42:318–37.
1978 "Social organization and innovation in rural Indian communities." American Rural Sociology Society Meetings, Madison, Wisconsin.

Gartrell, John W. and C. David Gartrell
1979 "Status, knowledge and innovation: risk and uncertainty in agrarian India." Rural Sociology 44:73–94.

Gotsch, Carl H.
1972 "Technological change and the distribution of income in rural areas." American Journal of Agricultural Economics 54:326–41.

Griffin, Keith
1974 The Political Economy of Agrarian Change: An Essay on the Green Revolution. London: MacMillan.
1978 International Inequality and National Poverty. New York: Holmes and Meier.

Hechter, Michael
1975 Internal Colonialism: The Celtic Fringe in British National Development, 1536–1966. London: Routledge and Kegan Paul.

Indian Institute of Public Opinion
1968 "A Portrait of Indian Poverty." Monthly Commentary on Indian Economic Conditions 10:45–62.

Iyengar, Kesava S.
1931 Economic Investigations in the Hyderabad State 1929–30 (five vol.). Hyderabad: Central Government Press.

Kravis, Irving B.
1962 The Structure of Income. Philadelphia: University of Pennsylvania Press.

Kuznets, Simon
1963 "Quantitative aspects of economic growth of nations, VII: the distribution of income by size." Economic Development and Cultural Change 11:1–80.

Lakshmanna, C.
1973 Caste Dynamics in Village India. Bombay: Mackiketa.

Lipton, Michael
1968 "A strategy for agriculture: urban bias and rural planning." Pp. 83–148 in Paul Streeten and Michael Lipton (eds.), The Crisis of Indian Planning. London: Oxford University Press.

Maddison, Angus
1971 Class Structure and Economic Growth. New York: Norton.

Mandelbaum, David C.
1970 Society in India (2 vol.). Berkeley: University of California Press.

Marriott, McKim
1965 Caste Ranking and Community Structure. Chicago: University of Chicago Press.

Mencher, Joan P.
1974 "Conflicts and contradictions in the 'Green Revolution': The Case of Tamil Nadu." Economic and Political Weekly 9:309–23.

Moore, Barrington
1966 Social Origins of Dictatorship and Democracy. Boston: Beacon.

Morrison, D. E., Krishna Kumar, Everett M. Rogers, and Frederick Fliegel
1976 "Stratification and risk taking: a further negative replication of Cancian's theory." American Sociological Review 41:912–9.

Myrdal, Gunnar
1968 Asian Drama: An Inquiry into the Poverty of Nations (3 vol.). New York: Random House.

Parthasarathy, G. and D. S. Prasad
1974 "Responses to, and impact of H.Y.V. rice according to land size and tenure in a delta village, Andhra Pradesh, India." The Developing Economies 12:182–98.

Paukert, Felix
1973 "Income distribution at different levels of development: a survey of evidence." International Labor Review 108:97–125.

Portes, Alejandro
1976 "On the sociology of national development: theories and issues." American Journal of Sociology 82:55–85.

Portes, Alejandro and John Walton
1976 Urban Latin America: The Political Condition from Above and Below. Austin: University of Texas Press.

Rao, Y. V. Lakshmana
1966 Communication and Development. Minneapolis: University of Minnesota Press.

Roy, Prodipto, Frederick Fliegel, Joseph Kivlin, and Lalit Sen
1968 Agricultural Innovation Among Indian Farmers. Hyderabad: National Institute of Community Development.

Rubinson, Richard
1976 "The world-economy and the distribution of income within states: a cross-national study." American Sociological Review 41:638–59.

Scott, James C.
1976 The Moral Economy of the Peasant: Rebellion and Subsistence in Southeast Asia. New Haven: Yale University Press.

Sharda, Bam Dev
1977 Status Attainment in Rural India. Delhi: Ajanta.

Sharma, Miriam
1978 The Politics of Inequality: Competition and Control in an Indian Village. Honolulu: University of Hawaii.

Singh, Vijai P.
1976 Caste, Class and Democracy: Changes in a Stratification System. Cambridge, Mass.: Schenkman.

Snipp, C. Matthew and Gene F. Summers
1980 "The welfare state in the community: a general model and empirical assessment." American Sociological Society Meetings, New York.

Stack, Steven
1978 "The effect of direct government involvement in the economy on the degree of income inequality: a cross-national study." American Journal of Sociology 84:1096–126.

Stanfield, David J. and Gordon C. Whiting
1972 "Economic strata and opportunity structure as determinants of innovativeness and productivity in rural Brazil." Rural Sociology 37:401–16.

Swensen, G. Geoffery
1976 "The distribution of benefits from increased rice production in Thanjavur District, South India." Indian Journal of Agricultural Economics 31:1–12.

Wallerstein, Immanuel
1974 The Modern World-System: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century. New York: Academic Press.

Weede, Erich
1980 "Beyond misspecification in sociological analysis of income inequality." American Sociological Review 45:497–501.

Weeks, John
1970 "Uncertainty, risk, and wealth and income distribution." Journal of Development Studies 6:28–36.

RACIAL DISCRIMINATION IN CRIMINAL SENTENCING: A CRITICAL EVALUATION OF THE EVIDENCE WITH ADDITIONAL EVIDENCE ON THE DEATH PENALTY*

GARY KLECK
Florida State University

Reevaluation of published research on racial bias in criminal sentencing and of data on execution rates by race from 1930 to 1967 and on death-sentencing rates from 1967 to 1978 indicates that, except in the South, black homicide offenders have been less likely than whites to receive a death sentence or be executed. For the 11% of executions imposed for rape, discrimination against black defendants who had raped white victims was substantial, but only in the South. Evidence for noncapital sentencing also largely contradicts a hypothesis of overt discrimination against black defendants. Although black offender–white victim crimes are generally punished more severely than crimes involving other racial combinations, this appears to be due to legally relevant factors related to such offenses. Crimes with black victims, however, are less likely than those with white victims to result in imposition of the death penalty. The devalued status of black crime victims is one of several hypothetical explanations of the more lenient sentencing of black defendants.

The legitimacy of the legal systems of modern democracies depends heavily on the degree to which the systems operate in a manner consistent with their own stated procedural standards of justice. It has been argued that Western societies, including the United States, are undergoing a "legitimation crisis" and that this is occurring specifically in criminal justice systems at least partly because they fail to live up to their stated commitments to treatment of defendants without regard to ascribed personal characteristics such as race, ethnicity, and gender or partially ascribed characteristics such as status or class position (Quinney, 1974; Chambliss and Seidman, 1971).[1]

It is widely believed, and frequently stated, that the criminal justice system has been in the past, and remains, racially discriminatory (e.g., Sutherland and Cressey, 1970; Clark, 1970). The most frequently cited category of evidence for this assertion has been research indicating more severe sentencing of black criminal defendants than white defendants, especially in imposition of the death penalty. As there have been at least sixty empirical studies of adult criminal sentencing published which refer to race, it is not surprising that at least one critic of the criminal justice system has asserted that evidence on racial discrimination in sentencing is probably the strongest evidence of racial bias in the criminal justice system (Overby, 1971:575). Because the outcomes of sentencing decisions are among the most visible of legal processing, the legal system's claim to legitimacy is especially dependent on the public's perception of the pattern of such outcomes. Therefore, it seems particularly important to take a close look at evidence bearing on this issue.

The first part of this paper attempts a comprehensive assessment of the published scholarly empirical research on racial bias in criminal sentencing in the U.S. in connection with both capital punishment and noncapital sentencing. One of the principal sources of distortion regarding this issue in the past has been selective citation of studies supporting one position or another; therefore, great care has been

* Direct all correspondence to Gary Kleck, School of Criminology, Florida State University, Tallahassee, Florida 32306.

[1] As used in this paper the word "race" is socially defined, referring to "a human group that defines itself and/or is defined by other groups as different from other groups by virtue of innate and immutable physical characteristics" (van den Berghe, 1967:9).

taken to be as exhaustive as possible in finding relevant studies. The second part of the paper presents new evidence on race and capital punishment, consisting of an analysis of execution rates for blacks and whites over the period 1930 to 1967, for the United States as a whole and for the South, and of death-sentencing rates for the period from 1967 to 1978.

The Varieties of Racial Bias: Some Conceptual Distinctions

At least five different practices can produce racial differentials in criminal sentences which are likely to be viewed as illegitimate or unjust.

(1) Overt racial discrimination against minority defendants. This refers to the imposition of more severe dispositions on members of a subordinate racial group, independent of their legally relevant individual merits, and primarily as a direct result of the conscious or unconscious racial prejudice of the sentencing decisionmakers.

(2) Disregard for minority crime victims. This would include the failure to sentence offenders (of any race) who victimize minority-group members as severely as those who victimize nonminority-group members.

(3) Class discrimination. This refers to more severe treatment of lower-class defendants as a consequence of class prejudice. It may be due to hostility or indifference of middle-class decisionmakers toward culturally different defendants, or because lower-class defendants better fit popular stereotypes of serious or dangerous criminals. Because blacks in the United States are disproportionately members of the lower class, class discrimination would affect them more heavily than whites, independent of any overt racial discrimination. This assumes that among the set of criminal defendants blacks are more likely to be lower class than whites, a debatable assumption considering the overwhelmingly lower-class character of criminal defendants of all races.

(4) Economic discrimination. When a society's legal system is structured so that significant private economic resources are required in order to effectively obtain full legal protection, this constitutes economic discrimination, even where there is no class discrimination (as defined above). If low-income defendants receive more severe sentences than middle-income defendants because they cannot afford to hire an outside private attorney or cannot make bail, this constitutes economic discrimination and could produce racial differentials in sentencing outcomes.

(5) Institutional Racism. This refers to the application, possibly in a universalistic fashion, of decisionmaking standards which in themselves have considerable consensual support (possibly even among minority members) but which result in less favorable outcomes for minority defendants. As used in the past, the term seems to have referred to, among other things, practices (2) through (4); but institutional racism in sentencing can take other forms as well. For example, if racial-minority defendants are likelier to have prior criminal convictions, then the use of prior record as a criterion for sentence determination will tend to produce less favorable outcomes for minority defendants and would therefore be an instance of institutional racism (for this view, see Burke and Turk, 1975, or Farrell and Swigert, 1978a). The establishment, by legislatures, of higher statutory penalties for crimes committed more frequently by racial-minority members than by others (such as violent interpersonal crimes) would also constitute institutional racism, regardless of the behavior of judges, prosecutors, and others who influence sentencing outcomes.

For the sake of verbal and conceptual clarity, it is misleading to label either (3) or (4) "racial" bias or discrimination. Although clearly unjust and certainly related to race in the United States, these practices are not directly racial in themselves, since they affect whites as well as blacks and could occur in jurisdictions or societies where no racial distinctions of any sort were made. Therefore, although reference will be made to them in examining the evidence regarding overt racial discrimination, these practices are not themselves primary objects of the analysis.

The concept of institutional racism is highly problematical. It is so flexible that any practice producing unfavorable sentencing outcomes for racial-minority members can be characterized as racist, no matter how the outcomes were produced, whether they were intentionally sought, and regardless of what criteria were involved in the decisions producing the outcomes. Any pattern of sentencing involving the crimes commonly dealt with by U.S. criminal courts could be construed as institutional racism, since blacks commit a disproportionate share of such crimes relative to their share of the population (see Hindelang, 1978, for a thorough discussion of the evidence for this statement) and therefore are bound to receive a disproportionate share of the criminal punishment, no matter how fairly the sentencing process is administered. Only an alteration of the social conditions producing differentials in the racial distribution of criminal behavior or a radical redefinition of which crimes the courts focus on could eliminate institutional racism of this sort. For these reasons, neither this study nor any other study of sentencing per se could reject the hypothesis of institutional racism in sentencing. Consequently, attention in this analysis will be primarily focused on overt racial discrimination and secondarily on disregard for minority crime victims.

ASSESSMENT OF PRIOR RESEARCH

This review is intended to be an exhaustive assessment of all scholarly empirical studies of race and criminal sentencing of adults in the United States published up through 1979. It does not cover the few studies relating ethnicity and sentencing, such as Castberg (1971) or Hall and Simkus (1975), nor studies of conviction, as opposed to sentencing, such as Forslund (1969). It does cover studies of the determination of degree of homicide of which defendants were convicted, such as Farrell and Swigert (1978a; 1978b), since such a determination is tantamount to determination of sentence. Also included are studies of commutations of death sentences, as these have commonly been cited as sentencing studies.

When two or more studies of the same data set, using very similar methods, have been published, only one is included, or the studies are treated as a single study (e.g., Wolfgang and Riedel, 1973 and 1975; Farrell and Swigert, 1978a and 1978b, and Swigert and Farrell, 1977; Lotz and Hewitt, 1977 and Hewitt, 1976). Studies of juvenile court dispositions are excluded since, properly speaking, juveniles are not sentenced to criminal penalties and determination of disposition is in any case substantially different from the adult sentencing process.[2] Further, the following kinds of studies are excluded: anecdotal or journalistic accounts, case studies, hypothetical-case sentencing studies (e.g., Johnston et al., 1973), purely theoretical studies, and other reviews of the literature. The review also ignores studies such as those of Sellin (1928, 1935) which simply compare sentence lengths of persons sentenced or executed, without any comparison to numbers arrested, convicted, etc. Finally, unpublished studies such as dissertations are excluded because the availability of such material is limited.[3]

Studies which fit the selection criteria were located through an iterative search process. An initial list of relevant studies was compiled through a search of Sociological Abstracts, Crime and Delinquency Abstracts, Legal Abstracts, and Hagan's (1974) bibliography. (About one-third of the studies reviewed here were also included in Hagan's review.)

---

[2] Liska and Tausig (1979) reviewed eight studies which examined race and juvenile court dispositions, noting that five of them show a significant race effect in the conventionally expected direction. However, only two of the five simultaneously controlled for offense seriousness and prior record. The crude state of research in this field is also indicated by the almost complete absence of attention to controlling for effects of juveniles' family circumstances or disposition. Although the philosophy of the juvenile court explicitly defines family stability as a legitimate factor to be considered, Liska and Tausig do not say a word suggesting that racial differences in family intactness could account for differences in juvenile court dispositions.

[3] However, for a sampling of recent doctoral dissertation on the subject the reader may consult Dison (1976), Hutner (1977), and San Marco (1979). A reading of their abstracts indicates findings highly congruent with the conclusion of my review.

Then the references listed in each of these studies were examined for further relevant studies, and so on, until no further leads were uncovered.

Capital Punishment Sentencing Studies

Table 1 summarizes in compact form the prior scholarly empirical research on racial discrimination and use of the death penalty, first regarding murder and then rape. The last column indicates whether racial differences were statistically significant and in the predicted direction. In a minority of studies my assessment of the evidence presented in a study differs from that of the authors. Two examples of this are noteworthy.

Bedau (1964) concluded that there was racial bias in the final disposition of persons sentenced to death, despite his own acknowledgment that his data showed no significant relationship between race and final disposition, and despite the fact that the observed relationship was in the opposite direction to that indicating discrimination against nonwhites. In New Jersey 66.2% of nonwhites sentenced to die were executed, compared to 68.4% of whites (Bedau, 1964:19). Further, Bedau claimed there was no significant relationship between race and execution when felony and nonfelony murders were separated, but reanalysis of his data reveals that in fact nonwhite murderers were significantly less likely to be executed for felony killings than were white felony murderers ($\chi^2 = 8.114$, 1 degree of freedom, $p = .01$). There was no race difference for nonfelony killings.

Bowers (1974:81–107) claimed to have found evidence which pointed "unmistakably to a pattern of racial discrimination in the administration of capital punishment in America" (p. 10). His conclusion was not limited to the South or to use of the death penalty for rape, although he laid particular stress on these areas. Regarding racial bias in use of the death penalty for murder, the evidence that Bowers produced was of two kinds: he argued first that the lower mean age at execution of nonwhites compared to whites was evidence of racial discrimination, and, second, that the lower percentage of cases appealed by nonwhites compared to whites indicated discrimination. While Bowers acknowledged that the age differentials at execution may simply reflect age differences at commission of the offense and arrest, he argued that the difference at arrest was not as large as that at execution. In fact, there was a 2.4 year difference in mean age at execution in Bowers' data (p. 80), while Wolfgang's (1958:70) homicide arrest data indicate an almost identical 2.3 year difference in the median age of black and white male homicide arrestees. Regarding the data on appeals, racial differentials in percentage receiving appeals of death sentences for murder since 1940 are largely confined to the South (see especially Bowers' Table 3–9), consistent with our interpretation of the findings of previous death-penalty discrimination studies. Nevertheless, Bowers concludes that, as indicated by age differentials at execution and differentials in percentage receiving higher appeals, "racial discrimination in northern and western states began to rival that in the South, at least for the period of decline in capital punishment" (1974:104).

Probably the most serious shortcoming of death-penalty discrimination studies is that they nearly all fail to control for prior criminal record. The one study which introduced such a control, Judson et al. (1969), found no evidence of racial discrimination, suggesting that apparent racial differences in other studies may actually have been due to racial differences in prior criminal activity. This hypothesis is supported by Wolfgang's (1958:175–6) findings that black homicide offenders were significantly more likely to have previous criminal arrest records than white homicide offenders. It is further supported by Hagan's (1974:366–8) reanalysis of Nagel (1969), which indicated that where crude controls for prior record (record/no record) were introduced, racial effects shrank. Where more adequate controls for prior record (number of prior convictions) were introduced, the racial difference disappeared altogether (Judson et al., 1969:1366–76), suggesting that the dichotomous measure of prior record may be inadequate for

Table 1. Empirical Studies of Racial Discrimination and the Death Penalty

Study	Jurisdiction(s)	Time Period Covered	Dependent Variables	Variables Controlled	Race Relationship Significant?
Murder:					
Mangum (1940)	9 Southern States	1909–1938	Executed/commuted	None	Yes
Johnson, G. (1941)	N. Carolina	1933–1939	Sentence; executed/commuted	None	Sentence, yes/no[a]; exec./commuted, no
Garfinkel (1949)	N. Carolina	1930–1940	Charge, conviction, sentence	Degree of Homicide	Yes/No[a]
Johnson, E. (1957)	N. Carolina	1909–1954	% executed, admissions to death row	None	Yes
Bensing & Schroeder (1960)	Cleveland	1947–1953	Death sentence/other	Degree of Homicide	No
Bridge & Mosure (1961)	Ohio	1950–1959	Executed/commuted	None	Yes
Wolfgang et al. (1962)	Pennsylvania	1914–1958	Executed/commuted	Felony/nonfelony, type of counsel[b]	Felony, yes; nonfelony, no
Bedau (1964)	New Jersey	1907–1960	Executed/commuted	Felony/nonfelony	No
Wolf (1964)	New Jersey	1938–1961	Death sentence/life imprisonment	Felony/nonfelony	No
Bedau (1965)	Oregon	1903–1964	Executed/commuted	None	No
Judson et al. (1969)	California	1958–1966	Death sentence/other sentence (imposed by penalty jury)	Prior record, occupation, characteristics of offense, and others	No
Bowers (1974)	U.S.	1864–1967	Appeal of death sentence	Region	In the South, yes, elsewhere, no
Rape:					
Johnson, O. [1951] (1970)	Louisiana	1900–1950	Executed/commuted	None	Yes
Johnson, E. (1957)	N. Carolina	1909–1954	% executed, admissions to death row	None	No
Florida Civil Liberties Union (1964)	Florida	1950–1964	Death sentence/other	None	Yes
Partington (1965)	Virginia	1908–1963	Death sentence/other	Type of rape	Yes
Wolfgang & Reidel (1973; 1975)	6 Southern States[c]	1945–1965	Death sentence/other	Contemporaneous offense/no contemporaneous offense	Yes

[a] There were no significant differences between black killers as a group and white killers as a group, but when offender/victim racial relationship was examined, blacks who killed whites were found to have received significantly more death sentences than the other three groups.

[b] Not controlled for simultaneously.

[c] Data were gathered on 11 states, but published results refer only to 6 states for some tables, 5 for other, 3 for yet other tables. Reidel has informed the author that the data for this study were partially destroyed in a fire, accounting for the fragmented data analysis. The 1975 study is simply an analysis of a subset of the data analyzed in the 1973 study; therefore, the two are treated as a single study.

control purposes (see also Green, 1961:11 on this point).

All of the studies purporting to find racial bias in use of the death penalty for murder failed to control for income, class, or occupation of the defendants.[4] However, the most methodologically sophisticated study of the subject, which did control for defendant's occupation, found no racial effect on whether or not a death sentence was imposed for murder by California penalty juries (Judson et al., 1969:1366–76). Further, they found no relationship between the victim/offender racial relationship and sentence imposed, suggesting that the findings of Johnson (1941) and Garfinkel (1949) may have reflected regional and temporal peculiarities characteristic of North Carolina (or more generally, the South) in the 1930s and earlier.

Several points should be noted about the pattern of findings on discrimination in use of the death penalty for murder. First, every single study consistently indicating discrimination towards blacks was based on older data from Southern states, and three of these four studies were based on overlapping data from North Carolina. Second, all of the studies finding discrimination in administration of capital punishment for murder were not in fact studies of sentencing, although most of them have been cited in the research literature as if they were. Mangum (1940), Johnson (1957), Wolfgang et al. (1962), and Bridge and Mosure (1961) all studied commutation of death sentences, not sentencing itself, while Bowers' (1974) data largely concerned appeals of death sentences. Third, all of these studies failed to control for prior criminal record of the defendant, for the defendant's class or income, or for the distinction between felony and nonfelony killings. Since studies which do introduce such controls find that they reduce the sentencing differentials between blacks and whites (Green, 1961; Judson et al., 1969; Nagel, 1969), even in the South the racial differential may have been due to differences in criminal record, income, or type of homicide committed rather than discrimination. The evidence considered as a whole indicates no racial discrimination in use of the death penalty for murder outside the South, and even for the South empirical support for the discrimination hypothesis is weak.

Regarding the use of capital punishment for rape, the evidence strongly suggests overt discrimination against black defendants. Four of the studies of this issue found evidence of discrimination, while the relationship between race and the carrying out of the death sentence (as opposed to commutation of the sentence) was not significant in the Johnson (1957) study, according to Hagan's (1964:370) reanalysis. The relationships found in the other four studies (Johnson, 1970; Florida Civil Liberties Union, 1964; Partington, 1965; Wolfgang and Reidel, 1973, 1975) are very strong, and the evidence indicates that the death penalty for rape was largely used for punishing blacks who had raped whites. Although all of these studies were methodologically crude, it is doubtful if additional controls could eliminate the huge racial differentials in use of the death penalty. The importance of this conclusion, however, is limited by several facts. First, the death penalty has rarely been used to punish rape. Only 11.8% of the executions from 1930 to the present were for rape. Virtually all the rest were for murder. Second, the use of the death penalty for rapes has always been, at least since national data on executions were first gathered in 1930, strictly a peculiarity of the South. Not a single execution for rape occurred outside the South or the border states during that period (U.S. Federal Bureau of Prisons, 1970). Third, the imposition of death sentences for rape has virtually disappeared. Of the 183 persons who were sentenced to death during 1978, only one was sentenced for rape (U.S. NCJISS, 1979:25). Thus, the rape discrimination conclusion is of historical significance with regard to capital punishment in the South, but has limited relevance to current debates over capital punishment, especially since the United

[4] Although both Wolfgang et al. (1962) and Johnson (1957) had data on occupation of offenders as well as race, they did not attempt to simultaneously control for race and occupation. No explanation of this conspicuous omission was offered in either study.

States Supreme Court declared, in Coker vs. Georgia (1977), that the use of the death penalty for the rape of an adult woman was a disproportionate penalty and therefore unconstitutional (U.S. NCJISS, 1979:2).

Noncapital Punishment Sentencing Studies

Table 2 summarizes the empirical research on sentencing involving penalties other than the death penalty. The results of each study are simply summarized in the last column of the table ("Discrimination?"), with each study characterized as to whether its findings largely supported a discrimination hypothesis (indicated by "Yes"), were mostly inconsistent with such a hypothesis ("No"), or were only partially consistent with the hypothesis ("Mixed"). Mixed findings most frequently occurred when more than one crime was studied, when sentencing patterns of a number of judges were reviewed, or when more than one measure of sentencing outcome was examined.

Studies were classified, in somewhat arbitrary fashion, according to what proportion of their findings were in favor of the discrimination hypothesis. They were characterized as mixed if from one-third to one-half (inclusive) of the findings favored the discrimination hypothesis and as favorable to the hypothesis if more than one-half of the findings favored it. For example, if a study examined eight different offenses, it would be labelled "Yes" if evidence of bias against black defendants was found for four or more offenses, as "Mixed" if such evidence was found for three of the offenses, and as "No" if two or fewer offenses showed such evidence. Since it could be argued that evidence of discrimination even for one crime or sentence-outcome measure out of many is evidence worth taking very seriously, readers must judge for themselves the significance of the "mixed" findings.[5]

Under the heading "Sentencing Measure" in Table 2, the dependent variable in each study is noted. The term "disposition" indicates that the dependent variable distinguished between categories like probation, jail sentence, prison sentence, etc., while "sentence severity" denotes a single scale of severity of disposition or sentence constructed by the researcher. The other terms are self-explanatory.

Table 2 also indicates whether the authors of these studies in any way controlled for the type of criminal offense involved. In those studies where only one type of crime was involved, or where several very similar offenses were studied, such a control was obviously unnecessary. However, where several different offense types were lumped together, as in Cargan and Coates (1974) or Pope (1975a), differences in sentence received by black and white defendants could be at least partly attributable to differences in the seriousness of the types of offenses for which they were prosecuted.

Of the 40 studies listed in Table 2, only eight consistently support the racial discrimination hypothesis, while 12 are mixed and the remaining 20 produced evidence consistently contrary to the hypothesis. Since a study's findings were characterized as mixed even if as few as a third of them favored the discrimination hypothesis, this means that a substantial majority of all of the findings of these 40 studies contradicted the hypothesis. However, the evidence for the hypothesis is even weaker than these numbers suggest, since of the minority of studies which produced findings apparently in support of the hypothesis, most either failed completely to control for prior criminal record of the defendant, or did so using the crudest possible measure of prior record—a simple dichotomy distinguishing defendants with some record from those without one. This is probably the most important flaw in studies drawing a conclusion of racial discrimination,

[5] Use of these standards occasionally resulted in characterizations of findings which differed from those of the original authors (e.g., Bedau, 1964). Since I attempted to accurately reflect the studies' theoretical or ideological preferences, when the original authors' conclusions did not seem congruent with their data, their conclusions were discounted. Skeptical readers are encouraged to examine the original studies in order to judge for themselves the accuracy of my characterizations.

Table 2. Race and Noncapital Sentencing

Study	Jurisdiction(s)	Time Period Covered	Sentencing Measure	Offenses Involved	Control for Offense Type?	Measure of Prior Record	Discrimination?
Martin (1935)	Texas	1930	Sentence	Various	Yes	None	Yes
Lemert & Rosberg (1948)	Los Angeles	1938	Prison sentence/other; sentence length	Various	Yes	No record/minor record/prior felony	Mixed
Bullock (1961)	Texas	1958 & before[a]	Prison sentence less than 10 yrs./over 10 yrs.	Burglary, rape, murder	Yes	None	Mixed
Green (1961)	Philadelphia	1956–57	Sentence	Burglary, robbery, theft misdemeanor	Yes	0–1 prior conviction/2+ prior convictions	No
Jacob (1962)	New Orleans	1954–60	Prison sentence greater than 1 yr/other	Various	Yes	None	Mixed
Cameron (1964)	Cook County (Chicago)	1943–50	% jail sentences; % 30 days or more	Shoplifting	†	None	Yes
Green (1964)	Philadelphia	1956–57	Sentence length	Robbery, burglary	Yes	Prior convictions	No
Howard (1967)	Baltimore	1962–66	Sentence length, type	Rape	†	None	Yes
Janos & Mendelsohn (1967)	Detroit	1966	% jailed, average fine	Traffic	No	None	No
Baab & Ferguson (1968)	27 counties in Texas	1966	Sentence severity	Various felonies	Yes	No. prior felony, misdemeanor convictions	No
Southern Regional Council (1969)	7 Southern states	1967 & before[a]	Sentence length	Various	Yes	1st offender/recidivist	Yes
Nagel (1969)	194 counties in 50 states; Federal	1962; 1963	Prison sentence/other; % prison sentence greater than 1 year	Assault, larceny	Yes	Prior conviction(s)/no prior convictions	Mixed
Atkinson & Newman (1970)	A major midwestern metropolitan center	c. 1968?	Fine, jail sentence/other	Nontraffic misdemeanors	No	None	No
Gerard & Terry (1970)	8 Missouri counties	1962	% prison/probation/fine	Various felonies	Yes	None	Yes
Mileski (1971)	"Middle-sized Eastern City"	c. 1969	Incarceration/other	Intoxication	†	Length and recency of prior record	No
Conklin (1972)	Boston	1964; 68	Disposition	Robbery	†	None	No
Levin (1972)	Minneapolis, Pittsburg	1966	Sentence length	Various	Yes	"prior record"	No
Rau (1972)	Federal	1967–70	% sentenced to prison, mean sentence length	Various	Yes	None	Mixed
Greenwood et al. (1973)	Los Angeles	1970	Sentence level (felony/misdemeanor)	Theft	†	None	No
Cargan and Coates (1974)	Montgomery County, Ohio	c. 1971–72	% sentenced to prison	Various felonies	No	None	Mixed

Burke & Turk (1975)	Indianapolis	1964	Disposition	Various	Yes	Prior arrest/none	No
Chiricos & Waldo (1975)	N. Carolina	1969–73	Sentence length	Various	Yes	Felony conviction/none, ≥5 arrests/<5 arrests; juvenile institutions/no	No
	S. Carolina	1969–70					
	Florida	1969–70					
Kulig (1975)	Douglas County, Nebraska	1970–72	Mean sentence, % receiving probation	Various	Yes	Low/med./high arrest record	Yes
Pope (1975a)	12 California counties	1969–71	Disposition	Various felonies	No	No arrests/some arrests; no prison/previous prison	Mixed
Pope (1975b)	12 California counties	1969–71	Disposition; sentence length	Burglary, assault	Yes	Same as above	No
Tiffany et al. (1975)	Federal	1967–68	Sentence length	Various	Yes	"Prior record"	Mixed
Clarke & Koch (1976)	Charlotte, N.C.	1971	Prison sentence/other	Burglary/larceny	Yes	Prior arrests	No
Kelly (1976)	Oklahoma	1974 & before[a]	Sentence length	Burglary, homicide	Yes	No prior convictions	Mixed
Rhodes (1976)	Minneapolis, St. Paul	1970	Prison sentence/other disposition	Various	Yes	Number & seriousness of prior convictions	Mixed
Zimring et al. (1976)	Philadelphia	1970	Death or life sentence/other prison	"Felony" murder	†	None	Yes
Bernstein et al. (1977)	City in N.Y. State	1974–75	Sentence disposition (7 levels)	Various	Yes	Weighted index of prior convictions	No
Lotz & Hewitt (1977)	King County, Washington	1973	Prison or jail/deferred, suspended sentence	Felonies	Yes	"Prior record"	No
Perry (1977)	U.S. Military	1972 & before	Sentence length	Various military offenses	Yes	None	No
Uhlman (1977)	"Major metropolitan center"	1968–74	Sentence length	Various	Yes	None	Yes
Farrell & Swigert (1978a; b)	"Large urban jurisdiction in N.E."	1955–73	Degree of homicide assigned	Homicides	†	None	No
Gibson (1978)	Atlanta	1968–70	% "severe sentences"	Various	Yes	Felony conviction/none	Mixed
Lizotte (1978)	Chicago	1971	Sentence length	Various	Yes	Arrest record/no record	Mixed
Sutton (1978)	Federal	1971	Disposition, sentence length	8 felonies	Yes	Conviction, incarceration	No
Foley & Rasche (1979)	Missouri	1959–74	Sentence length	Various felonies	Yes	None	No
Myers (1979)	Indianapolis	1974–76	Prison sentence/other sanction	Various felonies	Yes	Arrest/conviction/incarceration	No

[a] Sample of prisoners' records were examined in the year indicated, but the prisoners had been sentenced in that year and previous years.

† Only one offense type was studied, or offenses were so homogeneous that control for offense type or seriousness was unnecessary.

since the most methodologically sophisticated sentencing studies have consistently shown various measures of prior record to be either the strongest predictor, or among the strongest predictors, of sentences received (Chiricos and Waldo, 1975; Bernstein et al., 1977; Lotz and Hewitt, 1977; Lizotte, 1978). It appears to be the case that the more adequate the control for prior record, the less likely it is that a study will produce findings supporting a discrimination hypothesis.

Table 3 summarizes the whole body of prior research on race and sentencing, both capital and noncapital. Simply adding up the number of studies favoring or not favoring the discrimination hypothesis could be somewhat misleading, since some studies are clearly better than others and should therefore be weighted more heavily than others in assessing the body of evidence as a whole. Therefore, although it would be difficult to assign exact weights, some simple quality distinctions can be made, such as distinguishing between studies which control for prior criminal record and studies which do not. Regarding noncapital punishment, Table 3 makes clear the importance of such controls—one-third of the studies without a control for prior criminal record support a discrimination conclusion, while less than a tenth of those with such controls support a discrimination conclusion. Regarding capital punishment, separate tallies of studies with and without such controls are unnecessary, since only one study, that of Judson and his colleagues (1969), controlled for prior record, finding no evidence of racial discrimination either in the sentencing of black defendants in general or in sentencing of those who had victimized whites.

Table 3. Summary of Prior Research

	Results			
	Yes	Mixed	No	Total
Capital Sentencing				
All Studies	7	4	6	17
Murder	3	4	5	12
Rape	4	0	1	5
Noncapital Sentencing				
All Studies	8	12	20	40
Control for prior record	2	8	13	23
No control for prior record	6	4	7	17

NOTE: See the discussion of prior literature in the text for an explanation of the classification of studies by their results.

Interracial Relationship of Offender and Victim

It has long been argued that racial bias in sentencing is not to be detected only by looking at the race of the defendant, but by noting the racial relationship of the offender and the victim (e.g., Johnson, 1941). Specifically, it is asserted that crimes involving black offenders and white victims are punished more severely than crimes involving the other three racial combinations, either because crimes involving black victims are taken less seriously or because the crossing of racial lines in the commission of a crime is taken very seriously (Johnson, 1941; Garfinkel, 1949). While black offender–white victim crimes, especially homicides and rapes, are punished more severely than crimes with other racial combinations, it is unclear whether this is due to the racial character of the crime, or to related, confounding factors. Black offender–white victim killings are more likely than other killings to involve an offender and a victim who are strangers to each other, and such killings are much more severely punished regardless of the races involved (Lundsgaarde, 1977:232). Such killings are also more likely to be committed in connection with some other felony, like robbery. Data in Block and Zimring (1973:8) indicate that for Chicago homicides in 1970, 38% of killings with black offenders and white victims were robbery killings, while only 5% of the white offender–white victim killings were robbery killings. Felony killings are punished more severely than other homicides, regardless of races involved (Wolfgang et al., 1962; Bedau, 1964; Wolf, 1964). Finally, black-white killings are less likely than black-black killings to be victim-precipitated, and victim-precipitated killings in turn are less likely to be premeditated (Wolfgang, 1958), leading one to expect less severe punishment of black-black killings for this reason, rather than the racial relationship per se. Eleven

studies have examined sentencing outcomes by racial combination, and of these, seven (Johnson, 1941; Garfinkel, 1949; Florida Civil Liberties Union, 1964; Howard, 1967; Southern Regional Council, 1969; Wolfgang and Reidel, 1973; Zimring et al., 1976) found more severe punishment for black-white offenses.[6] However, none of these studies controlled for the possibly confounding factors we have mentioned. The only four studies which did introduce such controls (Green, 1964; Judson et al., 1969; Farrell and Swigert, 1978b; Myers, 1979) all found no evidence of such sentencing patterns. Thus, consideration of the pattern of findings as a whole strongly suggests that the interracial relationship itself does not affect the sentencing decision, except in connection with the punishment of rape in the South (Florida Civil Liberties Union, 1964; Howard, 1967; and Wolfgang and Reidel, 1973, 1975 support this limited assertion of discrimination).

Examination of prior studies on the question of racial discrimination and use of the death penalty for murder has suggested that many of their conclusions may be seriously time-bound and region-bound. Their findings may not be generalizable to areas outside the South, considering the generally contrary findings of studies of non-Southern jurisdictions using more recent data. Given these considerations, it would seem reasonable to study national sentencing practices, making regional comparisons, using data covering as long a period of time as possible.

EXECUTION RATES AND DEATH SENTENCING BY RACE

It has been claimed that "racial discrimination is strongly suggested by the national execution figures" (NAACP, 1971: 51–2). Clearly, blacks have been executed in numbers far out of proportion to their numbers in the population. Over the period 1930–1976, 53.6% of all legally executed persons in the United States were black, although blacks constituted only about 10–11% of the U.S. population during that period (U.S. Federal Bureau of Prisons, 1971:8; U.S. Bureau of the Census, 1977:25). This disproportion, however, cannot in itself be taken as evidence of racial discrimination, since blacks also commit a large proportion of U.S. homicides, the crime most frequently punished by death. A more meaningful measure of capital punishment sentencing outcome would be an indicator of execution risk, i.e., an execution rate. A true rate compares a number of events (such as executions) with the number of times the event could have occurred. Therefore, the ideal base for the execution rate could be the number of persons convicted of a capital offense, i.e., a crime for which, in a given jurisdiction, the offender could be sentenced to death. However, there are no national data on the number of such crimes committed or on persons arrested for the crimes. Therefore, a surrogate measure is needed.

In this analysis, execution risk by race is measured as the number of executions (for murder) of persons of a given race in a given year, divided by the number of homicide victims of that race who died in the previous year. The number of homicide victims of a given race is used as an approximation of the number of persons of that race who committed a homicide, whether a capital murder or a noncapital murder.[7] Since 92–97% of all homicides involve killers and victims of the same race (Garfinkel, 1949:371; Harlan, 1950:745; Wolfgang, 1958:379; Bensing and Schroeder, 1960:51; U.S. Federal Bureau of Investigation, 1977:9), the racial distribution of homicide victims can be used to describe the racial distribution

[6] Although Bullock (1961) has been cited in connection with the issue of sentencing and interracial relationships (e.g., Hindelang, 1969 and Baab and Ferguson, 1968), his study did not actually contain any data on victim-offender racial relationships.

[7] Execution rates for rape cannot be computed because there are no comparable data to use for the base of the rate. There were no national data on rape victimizations by race up until 1973 (by which time the judicial moratorium on execution had begun and even the imposition of death sentences for rape had virtually disappeared). In any case, the evidence showing discrimination in capital punishment of rape is fairly conclusive, making the computation of such rates redundant.

of homicide offenders with very little error (Wolfgang, 1958:223).

Since there is a median lag between arrest and a court trial for criminal homicide of slightly under six months (Wolfgang, 1958:296, 299), and a mean lag of about one year between conviction and execution (Lunden, 1962:1043; McCafferty, 1967:95; U.S. Federal Bureau of Prisons, 1970), the appropriate comparison for our purposes is between the executions in year t and arrests in year t−1, or possibly year t−2. For the sake of simplicity, the execution rates assume a one-year lag between commission of the homicide and execution of the offender over the time period studied. In any case, the results assuming a two-year lag would be substantively identical.

Table 4 presents the computed execution rates for blacks and whites, covering the entire period for which national execution data is available, 1930–1967.[8] In the final column, a ratio greater than one indicates a black execution rate higher than the white execution rate; therefore, for 25 of 38 of the years examined, the black execution rate was lower than the white execution rate. Since they are based on fairly small numbers of executions, race-specific execution rates and ratios of execution rates are somewhat unstable for single years, especially for the later years in the time series. Therefore the rates for the entire period were computed. For the period 1930–1967 there were 1,663 executions of whites for murder and 1,638 executions of blacks, while for the period 1929–1966 (lagged one year behind the other period) there were 159,482 white homicide victims and 168,518 black homicide victims (and presumably roughly equal numbers of homicide offenders). Therefore, the white execution rate for the entire period was 10.428 executions per 1,000 homicides and the black rate was 9.720 executions per 1,000 homicides. Thus, over the entire period, blacks were subject to a lower execution risk than whites.[9]

Given the regional pattern of discrimination findings of previous studies of capital punishment sentencing, it may be the case that execution rates are higher for blacks than for whites in the South and that this fact is obscured in national data. It is also possible that the relative execution risks of blacks and whites changes over time and by region. These possibilities are addressed using the data in Table 5.

These data indicate that the execution risk of black homicide offenders (actually nonwhites in this analysis) has indeed been greater than that of white homicide offenders in the South, while the opposite has been true in the rest of the United States. However, the excess of the black execution risk over the white execution risk in the South has declined over time, to the point where execution rates were roughly equal in the period since 1950. The evidence, considered in combination with prior research on capital punishment sentencing outcomes, suggests that use of the death penalty is not inevitably or inherently discriminatory, but rather that racial discrimination in its administration has been highly variable over time and between regions. These data support the racial discrimination hypothesis in connection with death penalty sentencing only for the South. Of particular interest is the somewhat surprising finding that in the recent past, outside of the South, the white execution risk has been substantially higher than the nonwhite risk, a fact which apparently has gone unnoticed in the literature. Possible explanations of this phenomenon will be discussed later in the paper.

[8] The number of homicide victims of each race excludes executions and killings committed by policemen in the line of duty. Executions are excluded because it is undesirable to have a common component in the numerator and denominator of the execution rate. Police killings are excluded because they are nearly always considered justifiable homicides and therefore not criminal. These exclusions make the homicide victim figures somewhat better surrogates for figures on criminal homicide offenders.

[9] It is debateable whether statistical tests of significance are appropriate where population data are involved, although Blalock (1972:238–9) has argued that they can serve to rule out an alternative explanation of a set of results—that the data could have been generated by chance processes rather than causal ones. A two-sample test of the difference between the proportions of persons executed among blacks and whites indicates the difference is significant at the .05 level (two-tailed test, Z=2.03).

Table 4. Execution Rates by Race, 1930–1967

Year	Black Executions for Murder	Black Homicide Victims[a]	White Executions for Murder	White Homicide Victims[a]	Black Execution Rate[b]	White Execution Rate[c]	Ratio of Black to White Execution Rates
1967	1	—	1	—	0.168	0.191	0.880
1966	0	5,945	1	5,230	0.000	0.205	0.000
1965	1	5,408	6	4,879	0.203	1.336	0.152
1964	4	4,926	5	4,492	0.893	1.197	0.748
1963	6	4,478	12	4,176	1.375	2.918	0.471
1962	15	4,364	26	4,112	3.583	6.468	0.554
1961	15	4,187	18	4.020	3.568	4.688	0.761
1960	26	4,204	18	3,840	6.394	4.826	1.325
1959	26	4,066	15	3,730	6.619	4.260	1.554
1958	20	3,928	20	3,521	5.040	6.073	0.839
1957	22	3,968	32	3,293	5.479	9.718	0.564
1956	31	4,015	20	3,239	8.105	6.240	1.299
1955	24	3,825	41	3,205	5.954	12.387	0.481
1954	33	4,031	37	3,310	8.317	11.315	0.735
1953	25	3,968	25	3,270	5.840	7.492	0.779
1952	36	4,281	35	3,337	9.217	10.965	0.841
1951	31	3,906	55	3,192	7.463	16.965	0.456
1950	32	4,154	36	3,362	7.402	9.882	0.749
1949	56	4,323	49	3,643	12.216	12.626	0.967
1948	61	4,584	32	3,880	13.475	8.095	1.664
1947	89	4,527	40	3,953	18.924	10.005	1.891
1946	61	4,703	45	3,998	15.877	12.879	1.233
1945	52	3,842	37	3,494	14.790	12.445	1.188
1944	48	3,516	45	2,973	13.829	14.227	0.972
1943	63	3,471	54	3,163	14.593	16.162	0.903
1942	58	4,317	57	3,341	13.075	16.681	0.784
1941	46	4,436	55	3,417	10.426	14.773	0.706
1940	61	4,412	44	3,723	13.610	11.429	1.191
1939	63	4,482	79	3,850	14.338	18.283	0.784
1938	63	4,394	89	4,321	13.011	20.408	0.638
1937	62	4,842	67	4,361	12.086	13.405	0.902
1936	93	5,130	86	4,998	18.383	15.826	1.162
1935	66	5,059	115	5,434	12.028	18.338	0.656
1934	89	5,487	64	6,271	17.056	9.672	1.763
1933	74	5,218	75	6,617	16.122	12.093	1.333
1932	63	4,590	62	6,202	13.322	10.003	1.332
1931	57	4,749	76	6,198	12.800	12.722	1.006
1930	57	4,453	90	5,974	13.106	16.474	0.796
1929	—	4,349	—	5,463	—	—	—

SOURCE: U.S., Federal Bureau of Prisons, *National Prisoner Statistics, Bulletin No. 46* (1971), p. 8; U.S., National Center for Health Statistics, *Vital Statistics of the U.S.: Mortality,* (annual issues, 1937–1966); U.S. Bureau of the Census, *Mortality Statistics,* (annual issues, 1929–1936).

NOTE: There were no executions in the U.S., 1968–1976; complete execution figures by race for the U.S. before 1930 are not available. Mortality figures for 1929–1932 refer to the death registration area rather than the entire U.S. (95.7% of the U.S. population was covered in 1929; 96.3% was covered in 1932). In 1929–1932 black homicide victim figures were estimated from "colored" homicide figures; the difference is very slight.

[a] Excluding executions and killings by policemen in the line of duty for 1950–1967. Figures before 1950 exclude executions but include killings by police.

[b] Black execution rate is number of black executions in year t per 1,000 black homicide victims in year t−1.

[c] White execution rate is number of white executions in year t per 1,000 white homicide victims in year t−1.

Possible Biases in Computation of Execution Risk

Our estimates of execution risk by race could be biased if the homicides which blacks commit are less likely to be capital murders than those committed by whites. If this were true, using the number of homicide victims of each race as the base of the execution rate would be misleading for comparative purposes, since a smaller proportion of the black offenders could be

Table 5. Execution Rates for Groups of Years by Region and Race

Years	White	Black[a]	Black Rate/ White Rate
		United States	
1930–1939	14.38	14.24	0.99
1940–1949	12.80	14.07	1.10
1950–1967	5.94	4.57	0.77
1930–1967	10.43	9.72	0.93
	White	Nonwhite[a]	Nonwhite Rate/ White Rate
		South	
1930–1939	11.01	14.41	1.31
1940–1949	11.24	14.26	1.27
1950–1967	5.02	5.30	1.06
1930–1967	8.39	10.47	1.25
		Non-South	
1930–1939	15.08	12.56	0.83
1940–1949	13.61	13.78	1.01
1950–1967	6.11	3.34	0.55
1930–1967	11.00	9.32	0.85

[a] In regional and state breakdowns of mortality by cause of death, the data refer only to white/nonwhite, while the published cumulations for the United States refer to white, black, and other races.

considered to be at risk of execution, compared to white offenders. In this case, the black execution would be understated relative to the white rate.

Three studies report figures on the percentage of criminal homicides designated as first degree (capital) murders, by race. One found the percentage designated first degree murder, both at indictment and at conviction, to be higher for blacks than for whites (Garfinkel, 1949:372), and one found the opposite (Bensing and Schroeder, 1960:43, 45, 88), while the third study found no significant difference (Wolfgang, 1958:302–03). Thus, no consistent relationship was found between race and proportion of criminal homicides designated first degree murder.

However, it has been argued that the designation of degree of homicide could itself be racially biased (Garfinkel, 1949). Would there be a racial difference if the degree of homicide were designated without bias? Given the hypothetical nature of the question, this is not easy to answer directly; however, we can evaluate it indirectly. It is generally agreed that there are a number of factors which can legitimately affect the designation of degree, including whether or not the homicide was committed in connection with another felony (called "felony killings"), whether the killing involved excessive violence or brutality, and of course whether or not the crime seemed to be premeditated. There is no direct evidence on premeditation by race. Wolfgang's (1958:376) data indicate that a higher proportion of killings committed by whites are committed in connection with robberies than are killings by blacks, suggesting that a higher proportion of white killings might be felony killings, compared to killings by blacks. However, the difference is slight, and this finding has not been corroborated elsewhere. On the other hand, Wolfgang (1958) found no significant relationship between race and the tendency to inflict multiple acts of violence, while black homicide arrestees were significantly more likely to have prior arrest records than white homicide arrestees (Wolfgang, 1958:160, 175–6). Therefore, there is little evidence that would indicate that killings committed by blacks are significantly less likely to be capital murders than those committed by whites.

There is another potential source of bias peculiar to the use of victim data by race as a surrogate for offender data by race. It was assumed that the number of offenders of one race would be roughly equal to the number of victims of that race. This assumption could be substantially incorrect, if, for example, killings involving black victims and white killers were more numerous than killings with white victims and black killers. If such were the case, the number of black victims would overestimate the number of black killers relative to white killers, and therefore underestimate the black execution rate relative to the white execution rate. Data relevant to this question are contained in the 1976 Uniform Crime Reports, which reported offender-victim racial relationships for murders and nonnegligent manslaughters (U.S. F.B.I., 1977). These data indicate that black offender–white victim killings are more numerous than white offender–black victim killings. They further indicate that while only 53.2% of the 10,538 homicide victims were black, 55.8% of the known offenders were black. This finding

suggests that use of victim data by race involves a bias whose correction would only strengthen our findings. Similar conclusions on victim-offender racial relationships could be drawn from data reported in smaller scale studies of criminal homicide (Garfinkel, 1949:371; Wolfgang, 1958:379; Bensing and Schroeder, 1960:51; and the seventeen-city study of Curtis, 1974:21).

Death-Sentencing Rates

Execution rates, as we have measured them, reflect not only rates at which defendants are sentenced to death, but also the extent to which such sentences are successfully appealed or commuted to a lesser penalty. Therefore, a purer measure of the rate at which defendants are sentenced to death is desirable. Annual data on the number of persons sentenced to die has been compiled, by race of the offender, for the United States since 1967. While this does not allow computation of death-sentencing rates for a very long period of time, it does update our analysis by providing information on the administration of capital punishment since the de facto moratorium on executions began in 1967.

Death-sentencing rates were first computed in a fashion similar to the computation of execution rates: the number of death sentences (actually, persons received by U.S. prisons from the courts, sentence of death) for murder is compared with the number of homicides in the previous year, for each race. Then a second measure of the death-sentencing rate was computed. It could be argued that a better measure of the risk of receiving a death sentence would use persons arrested for, or convicted of, capital crimes as the base for the rate, since it is only such persons who are actually at risk of receiving a death sentence. While there are no national data on convictions for murder, there are national figures on persons arrested for murder or nonnegligent manslaughter. Therefore, rates were computed using these figures for the base of the death-sentencing rate, again in a manner similar to the computation of execution rates.

Table 6 shows the computation of death-sentencing rates. The resulting rates, whether based on homicide deaths or homicide arrests, indicate that nonwhites were subject to a lower risk of being sentenced to death than whites over the period from 1967 to 1978. Because of the small numbers of death sentences each year, rates for single years are somewhat unstable, especially for 1972 and 1973. Nevertheless, the findings are on the whole quite consistent with the findings for execution rates.

This aggregate-level analysis does not show that there is never overt racial discrimination in the administration of the death penalty for murder outside of the South. There may be discrimination in particular jurisdictions, in specific individual cases, or at specific, previously unstudied, stages in the legal process leading up to execution, although a close reading of previous studies of various stages in this process, such as arrest, indicate that, at least outside the South, overt racial discrimination may be more apparent than real, just as seems to be the case with sentencing (regarding arrest, see Green, 1970; Black, 1971; Monahan, 1972; Lundman et al., 1978).

What the present analysis does show is that regardless of whatever discrimination there may or may not be at particular stages in the legal process, the outcome is a lower execution rate for blacks than for whites. If there is discrimination against blacks at one or more stages, then, given the observed net result, it seems that there must also be some compensating effects, favoring blacks, at other stages.[10] Likewise, if there is discrimination against blacks in one or more regions, jurisdictions, or specific cases, or with particular subtypes of homicides, then there must be some counterbalancing effects elsewhere.

The simple computation of execution and death-sentencing rates obviously does not in any way control for differences in prior criminal record (or other legally relevant variables, for that matter). Consid-

---

[10] This possibility of a pattern of compensating discriminatory effects has been raised by Nagel and Neef (1977:185–8).

Table 6. Death-Sentencing Rates by Race, 1967–1978.

Year	Nonwhite	White	Ratio: Nonwhite/White
	Sentences per 1,000 Homicide Deaths[a]		
1978	74/ 9,230 = 8.02	108/ 10,730 = 10.07	0.80
1977	64/ 9,439 = 6.78	68/ 10,115 = 6.72	1.01
1976	88/ 10,377 = 9.51	136/ 10,973 = 12.39	0.69
1975	143/ 10,817 = 13.22	121/ 10,648 = 11.36	1.16
1974	65/ 10,291 = 6.32	67/ 9,789 = 6.84	0.92
1973	23/ 10,498 = 2.19	11/ 8,840 = 1.24	1.76
1972	40/ 10,226 = 3.91	26/ 8,561 = 3.04	1.29
1971	51/ 9,045 = 5.64	45/ 7,803 = 5.77	0.98
1970	52/ 8,461 = 6.15	64/ 7,016 = 9.12	0.67
1969	36/ 7,880 = 4.57	49/ 6,806 = 7.20	0.63
1968	45/ 7,027 = 6.40	51/ 6,009 = 8.49	0.75
1967	38/ 6,077 = 6.25	36/ 5,230 = 6.88	0.91
1967–78	719/109,328 = 6.58	782/102,529 = 7.63	0.86
	Sentences per 1,000 Homicide Arrests[b]		
1978	74/ 9,256 = 7.99	108/ 7,866 = 13.73	0.58
1977	64/ 7,083 = 9.03	68/ 5,792 = 11.74	0.77
1976	88/ 8,592 = 10.24	136/ 6,581 = 20.67	0.50
1975	143/ 7,567 = 18.90	121/ 4,879 = 24.71	0.76
1974	65/ 7,677 = 8.74	67/ 5,236 = 12.80	0.66
1973	23/ 8,661 = 2.66	11/ 5,145 = 2.14	1.24
1972	40/ 8,586 = 4.66	26/ 4,716 = 5.51	0.85
1971	51/ 7,344 = 6.94	45/ 4,503 = 9.99	0.69
1970	52/ 6,669 = 7.79	64/ 3,743 = 17.10	0.46
1969	36/ 5,922 = 6.08	49/ 3,536 = 13.86	0.44
1968	45/ 5,018 = 8.97	51/ 3,200 = 15.94	0.56
1967	38/ 4,203 = 9.04	36/ 2,911 = 12.37	0.73
1967–78	719/ 86,578 = 8.30	782/ 58,126 = 13.45	0.62

Sources: Persons sentenced to death: U.S., NCJISS, *Capital Punishment* (1971–72; 1973; 1974; 1975; 1976; 1977; 1978).
Homicide deaths, 1966–76: U.S., NCHS, *Vital Statistics of the United States: Mortality (Year)* (1968–79).
Homicide deaths, 1977: U.S. NCHS, *Monthly Vital Statistics Report: Advance Report: Final Mortality Statistics 1977* (1979).
Arrests, 1966–77: U.S., F.B.I., *Crime in the United States (Year)* (1967–1978).

[a] Death sentences for murder, year t, per 1,000 homicide deaths, year t−1.
[b] Death sentences for murder, year t, per 1,000 homicide arrests, year t−1.

ering the stress laid on controlling for prior record earlier in the paper, this omission might seem to undercut confidence in the death penalty findings. However, because introduction of such controls has consistently reduced differences in sentencing outcome attributable to racial discrimination, correcting this omission would only tend to strengthen the conclusion of no overt discriminatory effect of homicide offenders' racial identity. In connection with capital punishment of rape, controlling for prior criminal record would reduce the apparent discriminatory effect to some degree, but this effect is so large to begin with that it is doubtful if the conclusion of discrimination would have to be altered.

Conclusions

The conclusions which can be drawn from the available evidence on the racial patterning of sentencing may be briefly summarized as follows:

(1) The death penalty has not generally been imposed for murder in a fashion discriminatory toward blacks, except in the South. Elsewhere, black homicide offenders have been less likely to receive a death sentence or be executed than whites.

(2) For the 11% of executions which have been imposed for rape, discrimination against black defendants who had raped white victims was substantial. Such discrimination was limited to the South and has disappeared because death sentences are no longer imposed for rape.

(3) Regarding noncapital sentencing, the evidence is largely contrary to a hypothesis of general or widespread overt discrimination against black defendants, although there is evidence of discrimination for a minority of specific jurisdictions, judges, crime types, etc.

(4) Although black offender–white victim crimes are generally punished more severely than crimes involving other racial combinations, the evidence indicates that this is due to legally relevant factors related to such offenses, not the racial combination itself.

(5) There appears to be a general pattern of less severe punishment of crimes with black victims than those with white victims, especially in connection with imposition of the death penalty. In connection with noncapital sentencing, the evidence is too sparse to draw any firm conclusions.

None of these findings are inconsistent with the assertion of institutional racism or income discrimination in sentencing. It is quite possible that low income makes it more difficult to make bail, hire a private attorney genuinely independent of the court, etc., for both blacks and whites, and that these factors in turn result in more severe sentencing outcomes, as Lizotte's (1978) research indicates.[11] If black criminal defendants are poorer than white criminal defendants, then income discrimination would produce racial differentials in sentences received. Nor are the data inconsistent with a hypothesis of overt discrimination at earlier stages of the criminal justice process. We might expect violations of stated values such as equal protection and justice for all to occur most commonly in connection with the least visible decisions, such as the decision to arrest, charge, prosecute, or release a defendant on bail (e.g., see Hagan, 1975 on the decision to charge). However, these decisions are less well studied than the sentencing decision, so the evidence for discrimination is necessarily even weaker than that regarding sentencing, quite apart from the actual prevalence of discriminatory practice.

The findings of this study do not suggest a different explanation for a well-known phenomenon. Rather they point to a phenomenon to be explained which differs from that conventionally addressed by American students of the legal reaction to crime and criminals. Students of the criminal justice system, concerned with the contemporary consequences of a historical pattern of racism, have sought to explain patterns of more severe treatment of blacks, while overlooking or downplaying the pattern of more lenient treatment of black defendants.

Blacks in the United States, both in the recent and more remote past, have been less likely than whites to receive a death sentence if they committed a homicide. Furthermore, this pattern is apparently not entirely limited to the sentencing of capital offenders. For a variety of specific crimes, jurisdictions, and judges, various researchers have produced data indicating more lenient treatment of black defendants than whites, although the admittedly scattered findings were usually deemphasized or discounted as merely anamalous results attributable to some flaw in the analysis or research design.[12] For example, Bullock (1961) found significantly shorter prison sentences were assigned to blacks convicted of murder; Levin's (1972) Pittsburgh data indicate that blacks received more lenient dispositions than whites for eight out of nine offense categories; and Bernstein and her colleagues (1977) found that blacks received significantly less severe sentences than whites. Gibson (1978:469) studied sentences given by individual judges and found that seven of eleven judges gave a higher percentage of severe sentences to whites than to blacks.

[11] It is interesting that Lizotte's path-analysis findings indicate that defendant's race affects whether the defendant makes bail (which in turn affects sentence), but that it does not affect the bail amount set. This suggests that there is no overt racial discrimination in bail setting, but that there is income discrimination. Since Lizotte had no measure of defendant's income, it is possible that the race effect was found simply because the race variable was serving as a rough surrogate for defendant income.

[12] This may be one of the more important subsidiary findings of the literature review. It is a chronic problem in this area, and perhaps in sociology as a whole, that researchers fail to recognize the significance of anomalies, which should alert them to the possible need for alterations in their fundamental assumptions rather than just their methods (See Kuhn, 1962: ch. VI on this issue).

The specification of phenomena to be explained is in a way a more fundamental scientific task than the development of explanations, since the former obviously must occur before the latter can even be imagined. The pattern of lenient treatment of black defendants in the South was recognized in the 1940s and before by observers such as Dollard (1937) and Myrdal ([1944] 1972), and various explanations were developed to account for it. Today, however, this phenomenon is largely disregarded.

A number of factors which may help account for this pattern can be briefly outlined.

(1) Blacks as devalued crime victims. Perhaps the most plausible explanation of lenient treatment of black offenders who commit predominantly intraracial crimes such as homicides, assaults, and rapes is that crimes with black victims are considered by predominantly white social-control agents to be less serious offenses, representing less loss or threat to the community than crimes with white victims (Myrdal, [1944], 1972:551). Thus, paradoxically, racist sentiments would produce more favorable treatment for members of the subordinate racial group who commit intraracial crimes.

(2) White paternalism. Students of criminal justice in the South have suggested a widespread view among whites of blacks as child-like creatures who were not as responsible for their actions as whites were, and who therefore could not be held accountable to the law to the extent that whites are (Dollard, 1937; Myrdal, [1944] 1972; Garfinkel, 1949). Therefore their perceived diminished responsibility presumably earned them more lenient sentences.

(3) Sociology-based tolerance. White paternalism may account for Southern sentencing patterns of the past but not patterns in the rest of the country in more recent times. However, it may have been replaced by a new form of white tolerance for black crime, involving the following line of reasoning: "Blacks commit crimes because of poverty, racism, and/or the resulting black poverty-subculture, which accepts or encourages criminal behavior. Their crimes are due to forces beyond their control or at least are to be expected in this light. Therefore blacks should not be held as responsible for their actions as whites." Of course, this is largely speculative; however, criminal court informants questioned by Bernstein and her colleagues (1977:753) stated that "some judges and prosecutors assume that nonwhites commit crimes because the nonwhite subculture accepts such behavior. These subcultural differences are considered by the judges and prosecutors, thereby making the offenses of nonwhites seem less pernicious."

(4) Affirmative action in the courts. White guilt over acknowledged past discrimination could motivate liberal criminal-justice decisionmakers to consciously or unconsciously compensate with more lenient treatment of black defendants.

(5) Compensation for institutional racism. Recognizing the handicaps of low income and greater prior criminal records which black defendants bring into court, some decisionmakers may attempt to compensate in determining sentence.

(6) Compensation for unconscious prejudice. Johnston et al. (1973) conducted a study involving criminal court judges sentencing hypothetical defendants and found that the hypothetical white "defendants" were sentenced more severely than the black "defendants." Two of the judges who participated in the study explained that they consciously sought to compensate in their sentencing for any unconscious prejudice on their part against minorities (p. 870).

Various combinations of these explanations can be used to account for less severe sentencing of blacks when and where it has occurred. Factors (1) and (2) may be primarily responsible for lenient sentencing in intraracial cases in the South in the 1940s and before, while different combinations of factors (3) through (6) account for leniency when it occurs elsewhere today and in the relatively recent past—say, since the late 1960s. Only future research designed to test the hypotheses can determine which of these are more than merely plausible and actually produce the patterns observed.

The findings of this paper should not be

interpreted as being incompatible with conflict, critical, or Marxist approaches to law, but rather only with the more simplistic, instrumentalist versions of these perspectives, which heavily stress the failure of the criminal justice system to operate according to its own stated standards of equity and proper procedure (e.g. Quinney, 1974; Chambliss and Seidman, 1971). In contrast, Beirne (1979) has pointed out the partial autonomy from particular social classes which the legal system enjoys, and has stated that "The capitalist class *as a whole* cannot be well served by frequent and visible abuses of due process" (p. 379). The criminal justice system can routinely operate to further legitimate the existing order through an obedience to its own rules and limits to its power, even when overt class or race bias in specific situations would otherwise aid particular segments of the ruling class. Clearly then, a more intellectually mature version of conflict theory would not necessarily predict overt class or race bias in the allocation of penalties among criminal defendants.

However, there are forms of class bias in the legal system which are not so clearly a threat to the legitimacy of the system. For example, the American legal system openly permits differing economic resources to be used in mounting a criminal defense, and such differences render legal advantages in avoiding conviction or obtaining lenient sentences if convicted, even though the advantages may operate indirectly and may involve no intentional prejudice on the part of any system decisionmaker. Lizotte (1978) has demonstrated how criminal sentence is affected by whether the defendant made bail and by the type of attorney the defendant had (private attorneys who were not courtroom regulars were more successful in negotiating light sentences than other types). While Lizotte had no measure of income, these are both clearly advantages more available to defendents with greater income. If equal protection of the law is a commodity which must be purchased, then this "equality" cannot be anything more than a legal fiction as long as the resources for such a purchase are distributed in an unequal fashion.

Serious though this economic or income discrimination in court processing may be, there is a far more fundamental bias in criminal sentencing. No studies of court processing of criminal defendants can address the issue of how legislatures criminalize behaviors common to lower-class persons, while either failing to criminalize or assigning slight penalties to equally harmful behaviors common among middle- or upper-class persons, such as poisoning of the air and water, manufacture of food, drugs, and other products harmful to human health, price-fixing, and consumer fraud. Detailed study of the use of wealth and power in controlling the ideological composition of legislatures and enforcement agencies, thereby influencing selection of behaviors to be criminalized, the original setting of penalty ranges, the determination of enforcement priorities, and allocation of enforcement resources, is likely to reveal far more about why blacks and lower-class persons are overrepresented in arrest, court, and prison data than studies of processing within the criminal justice system. The focus on the influence of ascribed characteristics of individual criminal defendants on processing decisions has, at least up to now, failed to yield the empirical support which would justify the attention that continues to be lavished on the subject.

REFERENCES

Atkinson, David N. and Dale A. Newman
1970 "Judicial attitudes and defendant attributes: some consequences for municipal court decisionmaking." Journal of Public Law 19:68–87.

Baab, G. A. and W. R. Ferguson
1968 "Texas sentencing practices: a statistical study." Texas Law Review 45:471–503.

Bedau, Hugo Adam
1964 "Death sentences in New Jersey." Rutgers Law Review 19:1–55.
1965 "Capital punishment in Oregon 1903–1964." Oregon Law Review 45:1–39.
1967 "Criminal justice: the general aspects." Pp. 405–14 in Hugo Adam Bedau (ed.), The Death Penalty in America. Garden City, N.Y.: Doubleday.

Beirne, Piers
1979 "Empiricism and the critique of Marxism on law and crime." Social Problems 26:373–85.

Bensing, Robert C. and Oliver J. Schroeder
1960 Homicide in an Urban Community. Springfield, Ill.: Charles Thomas.

Bernstein, Ilene Nagel, William R. Kelly, and Patricia A. Doyle
1977 "Societal reaction to deviants: the case of criminal defendants." American Sociological Review 42:743–55.

Black, Donald
1971 "The social organization of arrest." Stanford Law Review 23:1087–111.

Blalock, Hubert A.
1972 Social Statistics. New York: McGraw-Hill.

Block, Richard and Franklin E. Zimring
1973 "Homicide in Chicago, 1965–1970." Journal of Research in Crime and Delinquency 10:1–12.

Bowers, William J.
1974 Executions in America. Lexington, Mass.: D.C. Heath.

Bridge, Franklin M. and Jeanne Mosure[1]
1961 Capital Punishment. Staff Research Report No. 26. Columbus, Ohio: Ohio Legislature Service Commission.

Bullock, H. A.
1961 "Significance of the racial factor in the length of prison sentences." Journal of Criminal Law, Criminology and Police Science 52:411–7.

Burke, Peter and Austin Turk
1975 "Factors affecting postarrest dispositions: a model for analysis." Social Problems 22:313–32.

Cameron, Mary Owen
1964 The Booster and the Snitch. New York: Free Press.

Cargan, Leonard and Mary A. Coates
1974 "Indeterminate sentence and judicial bias." Crime and Delinquency 20:144–56.

Castberg, A. Didrick
1971 "The ethnic factor in criminal sentencing." Western Political Quarterly 24:425–37.

Chambliss, William J. and Robert B. Seidman
1971 Law, Order, and Power. Reading, Mass.: Addison-Wesley.

Chiricos, Theodore G. and Gordon P. Waldo
1975 "Socioeconomic status and criminal sentencing: an empirical assessment of a conflict proposition." American Sociological Review 40:753–72.

Clark, Ramsey
1970 Crime in America: Observations on Its Nature, Causes, Prevention and Control. New York: Simon and Schuster.

Clark, Stevens H. and Gary G. Koch
1976 "The influence of income and other factors on whether criminal defendants go to prison." Law and Society Review 11:57–92.

Conklin, John E.
1972 Robbery and the Criminal Justice System. Philadelphia: Lippencott.

Curtis, Lynn A.
1974 Criminal Violence. Lexington, Mass.: D.C. Heath.

Dison, Jack Everitt
1976 "An empirical examination of conflict theory: race and sentence length." Ph.D. dissertation, Department of Sociology, North Texas State University, Denton, Texas.

Dollard, John
1937 Caste and Class in Southern Town. New Haven: Yale University Press.

Farrell, Ronald A. and Victoria Lynn Swigert
1978a "Prior offense as a self-fulfilling prophecy." Law and Society Review 12:437–53.
1978b "Legal disposition of inter-group and intra-group homicides." Sociological Quarterly 19:565–76.

Florida Civil Liberties Union
1964 Rape: Selective Electrocution Based on Race. Miami: Florida Civil Liberties Union.

Foley, Linda A. and Christine E. Rasche
1979 "The effect of race on sentence, actual time served and final disposition of female offenders." Pp. 93–106 in John A. Conley (ed.), Theory and Research in Criminal Justice: Current Perspectives. Cincinnati: Anderson.

Forslund, Morris A.
1969 "Age, occupation and conviction rates of white and negro males: a case study." Rocky Mountain Social Science Journal 6:141–6.

Garfinkel, Harold
1949 "Research note on inter- and intra-racial homicides." Social Forces 27:369–81.

Gerard, Jules and T. R. Terry
1970 "Discrimination against negroes in the administration of criminal law in Missouri." Washington University Law Quarterly 1970:415–37.

Gibson, James L.
1978 "Race as a determinant of criminal sentences: a methodological critique and a case study." Law and Society Review 12:455–78.

Green, Edward
1961 Judicial Attitudes in Sentencing. London: MacMillan.
1964 "Inter- and intra-racial crime relative to sentencing." Journal of Criminal Law, Criminology and Police Science 55:348–58.
1970 "Race, social status, and criminal arrest." American Sociological Review 35:476–90.

Greenwood, Peter W., Sorrel Wildhorn, Eugene C. Poggio, Michael J. Strumwasser, and Peter De Leon
1973 Prosecution of Adult Felony Defendants in L.A. County: A Policy Perspective. Law Enforcement Assistance Administration, Washington, D.C.: U.S. Government Printing Office.

Hagan, John
1974 "Extra-legal attributes and criminal sentencing: an assessment of a sociological viewpoint." Law and Society Review 8:357–83.
1975 "Parameters of criminal prosecution: an application of path analysis to a problem of criminal justice." Journal of Criminal Law and Criminology 65:536–44.

Hall, E. L. and A. A. Simkus
1975 "Inequality in the type of sentences received by native Americans and whites." Criminology 13:199–222.

Harlan, Howard
1950 "Five hundred homicides." Journal of Criminal Law and Criminology 40:736–52.
Hewitt, John D.
1976 "Individual resources, societal reaction, and sentencing disparity." Western Sociological Review 7:31–56.
Hindelang, Michael J.
1969 "Equality under the law." Journal of Criminal Law, Criminology and Police Science 60:306–13.
1978 "Race and involvement in crimes." American Sociological Review 43:93–109.
Howard, Joseph C., Jr.
1967 Rape death penalty study reported in the New York Times, September 18, 1967, p. 33.
Hutner, Michael
1977 "Sentencing in Massachusetts Superior Court." Ph.D. dissertation, Department of Sociology, Syracuse University, Syracuse, N.Y.
Jacob, Herbert
1962 "Politics and criminal prosecutions in New Orleans." Tulane Studies in Political Science 8:77–98.
Janos, Dean and Robert I. Mendelsohn
1967 "The judicial role and sentencing behavior." Midwest Journal of Political Science 11:471–88.
Johnson, Guy
1941 "The negro and crime." The Annals of the American Academy of Political and Social Science 217:93–104.
Johnson, Elmer H.
1957 "Selective forces in capital punishment." Social Forces 36:165–9.
Johnson, Oakley C.
[1951] "Is the punishment of rape equally administered to negroes and whites in the
1970 state of Louisiana?" Pp. 216–28 in William L. Patterson, (ed.), We Charge Genocide. New York: International Publishers.
Johnston, Barbara L., Nicholas P. Miller, Ronald Schoenberg, and Laurence Ross Weatherly
1973 "Discrimination in felony sentencing—a study of influencing factors." Washington Law Review 48:857–89.
Judson, Charles J., James J. Pandell, Jack B. Owens, James L. McIntosh, and Dale L. Matschullat
1969 "A study of the penalty jury in first degree murder cases." Stanford Law Review 21:1297–431.
Kelly, Henry E.
1976 "Comparison of defense strategy and race as influences in differential sentencing." Criminology 14:241–9.
Kuhn, Thomas S.
1962 The Structure of Scientific Revolutions. Chicago: University of Chicago Press.
Kulig, Frank
1975 "Plea bargaining, probation, and other aspects of conviction and sentencing." Creighton Law Review 8:938–54.
Lemert, Edwin M. and Judy Rosberg
1948 The Administration of Justice to Minority Groups in L.A. County. Berkeley: University of California Press.
Levin, Martin A.
1972 "Urban politics and judicial behavior." Journal of Legal Studies 1:220–1.
Liska, Allen E. and Mark Tausig
1979 "Theoretical interpretations of social class and racial differentials in legal decisionmaking for juveniles." The Sociological Quarterly 20:197–207.
Lizotte, Alan J.
1978 "Extra-legal factors in Chicago's criminal courts: testing the conflict model of criminal justice." Social Problems 25:564–80.
Lotz, Roy and John D. Hewitt
1977 "The influence of legally irrelevant factors on felony sentencing." Sociological Inquiry 47:39–48.
Lunden, Walter A.
1962 "Time lapse between sentence and execution: the U.S. and Canada compared." ABA Journal 38:1043–5.
Lundsgaarde, Henry P.
1977 Murder in Space City: A Cultural Analysis of Houston Homicide Patterns. New York: Oxford University Press.
Lundman, Richard J., Richard E. Sykes, and John P. Clark
1978 "Police control of juveniles: a replication." Journal of Research in Crime and Delinquency 15:74–91.
Mangum, Charles S., Jr.
1940 The Legal Status of the Negro. Chapel Hill: North Carolina Press.
Martin, Roscoe
1934 The Defendant and Criminal Justice. University of Texas Bulletin No. 3437. Austin, Texas. Bureau of Research in the Social Sciences.
McCafferty, James A.
1967 "The death sentence, 1960." Pp. 90–103 in Hugo Adam Bedau, (ed.), The Death Penalty in America. Garden City, N.Y.: Doubleday.
Mileski, Maureen
1971 "Courtroom encounters: an observation of a lower criminal court." Law and Society Review 5:473–538.
Mohahan, T. P.
1972 "The disposition of juvenile offenders by race and sex in relation to the race and sex of police officers." The International Review of Modern Sociology 2:91–101.
Myers, Martha A.
1979 "Offended parties and official reactions: victims and the sentencing of criminal defendents." Sociological Quarterly 20:529–40.
Myrdal, Gunnar
[1944] An American Dilemma. New York: Pan-
1972 theon.
National Association for the Advancement of Colored People (NAACP)
1971 "Legal defense fund brief for Aikens 1971." Cited on p. 71 in William J. Bowers (ed.), Executions in America. Lexington, Mass.: D.C. Heath.

Nagel, Stuart
1969 The Legal Process From a Behavioral Perspective. Homewood, Ill.: Dorsey Press.
Nagel, Stuart and Marian Neef
1977 The Legal Process: Modeling the System. Beverly Hills: Sage.
Overby, Andrew
1971 "Discrimination against minority groups." Pp. 569–81 in Leon Radzinowicz and Marvin Wolfgang (eds.), Crime and Justice, Vol. II: The Criminal in the Arms of the Law. New York: Basic Books.
Partington, Donald
1965 "The incidence of the death penalty for rape in Virginia." Washington and Lee Law Review 22:43–75.
Perry, R. W.
1977 "Justice system and sentencing: the importance of race in the military." Criminology 15:225–34.
Pope, Carl E.
1975a Sentencing of California Felony Offenders. National Criminal Justice Information and Statistics Service, Washington, D.C.: U.S. Government Printing Office.
1975b The Judicial Processing of Assault and Burglary Offenders in Selected California Counties. National Criminal Justice Information and Statistics Service, Washington, D.C.: U.S. Government Printing Office.
1977 "The influence of social and legal factors on sentence dispositions: a preliminary analysis of offender-based transaction statistics." Journal of Criminal Justice 5:203–41.
Quinney, Richard
1974 Critique of Legal Order. Boston: Little, Brown.
Rau, Richard M.
1972 Sentencing in the Federal District Courts. Law Enforcement Assistance Administration, Washington, D.C.: U.S. Government Printing Office.
Rhodes, William M.
1976 "The economics of criminal courts: a theoretical and empirical investigation." Journal of Legal Studies 5:311–40.
San Marco, Louis Robert
1979 "Differential sentencing patterns among criminal homicide offenders in Harris County, Texas." Ph.D. dissertation, Department of Sociology, Sam Houston State University, Huntsville, Texas.
Sellin, Thorsten
1928 "The negro criminal: a statistical note." The Annals 140:52–64.
1935 "Race prejudice in the administration of justice." American Journal of Sociology 41:212–17.
Southern Regional Council
1969 Race Makes the Difference. Atlanta: Southern Regional Council.
Sutherland, Edwin H. and Donald R. Cressey
1970 Principles of Criminology. Philadelphia: Lippincott.
Sutton, Paul
1978 Variations in Federal Criminal Sentences: A Statistical Assessment at the National level. National Criminal Justice Information and Statistics Service, Washington, D.C.: U.S. Government Printing Office.
Swigert, Victoria Lynn and Ronald A. Farrell
1977 "Normal homicides and the law." American Sociological Review 42:16–32.
Tiffany, Lawrence, Yakov Avichai, and Geoffrey Peters
1975 "A statistical analysis of sentencing in Federal courts: defendants convicted after trial, 1967–1968." Journal of Legal Studies 4:369–90.
Uhlman, Thomas H.
1977 "The impact of defendant race in trial-court sanctioning decisions." Pp. 19–51 in John A. Gardiner (ed.), Public Law and Public Policy. New York: Praeger.
U.S. Bureau of the Census
1929–1936 Mortality Statistics. Washington, D.C.: U.S. Government Printing Office.
1975 Historical Statistics in the United States, Colonial Time to 1970. Washington, D.C.: U.S. Government Printing Office.
1977 Statistical Abstract of the United States. Washington, D.C.: U.S. Government Printing Office.
U.S. Federal Bureau of Investigation
1954–1978 Crime in the United States. Washington, D.C.: U.S. Government Printing Office.
U.S. Federal Bureau of Prisons
1970 National Prisoner Statistics, Bulletin No. 47, Capital Punishment, 1930–1970. Washington, D.C.: U.S. Government Printing Office.
1971 National Prisoner Statistics, Bulletin No. 46. Washington, D.C.: U.S. Government Printing Office.
1976 National Prisoner Statistics, Capital Punishment, 1975. Washington, D.C.: U.S. Government Printing Office.
U.S. National Center for Health Statistics
1937–1979 Vital Statistics of the United States (year), Mortality, Vol. I, Vol. II, Part A. Washington, D.C.: U.S. Government Printing Office.
U.S. National Criminal Justice Information and Statistics Service (NCJISS)
1971–1979 Capital Punishment. Washington, D.C.: U.S. Government Printing Office.
1976 Criminal Victimization in the United States: 1973. Washington, D.C.: U.S. Government Printing Office.
van den Berghe, Pierre L.
1967 Race and Racism: A Comparative Perspective. New York: Wiley.
Wolf, Edwin D.
1964 "Abstract of analysis of jury sentencing in capital case." Rutgers Law Review 19:56–64.
Wolfgang, Marvin E.
1958 Patterns in Criminal Homicide. Philadelphia: University of Pennsylvania Press.
Wolfgang, Marvin E., Arlene Kelly, and Hans C. Nolde
1962 "Comparison of the executed and com-

muted among admissions to death row." Journal of Criminal Law, Criminology, and Police Science 53:301–11.
Wolfgang, Marvin E. and Marc Riedel
1973 "Race, judicial discretion, and the death penalty." The Annals of the American Academy of Political and Social Science 407:119–33.
1975 "Rape, race, and the death penalty in Georgia." American Journal of Orthopsychiatry 45:658–68.
Zimring, Franklin E., Joel Eigen, and Sheila O'Malley
1976 "Punishing homicide in Philadelphia: perspectives on the death penalty." University of Chicago Law Review 43:227–52.

A TEST OF THE STABILITY OF PUNISHMENT HYPOTHESIS: THE CASE OF CALIFORNIA, 1851–1970*

RICHARD A. BERK
DAVID RAUMA
University of California at Santa Barbara
SHELDON L. MESSINGER
University of California at Berkeley
THOMAS F. COOLEY
University of California at Santa Barbara

Building on the works of Emile Durkheim, a number of scholars have claimed that, for a given society over time and in the absence of major societal upheavals, the proportion of people subjected to punishment by the state closely approximates a constant. In this paper, we address the "Stability of Punishment Hypothesis" with data from California for the period from the opening of the prison system in 1851 to 1970. The Stability of Punishment Hypothesis is formulated through recent modeling developments in macroeconomics, with which the presence of equilibrating tendencies can be explicitly represented, and for the first time direct tests are undertaken. No evidence for the hypothesis is found.

INTRODUCTION

It has often been asserted that societies are characterized by stable levels of punishment. Over relatively long periods and in the absence of social upheavals, the number of people punished should represent some constant fraction of the population. Across societies that fraction may vary, but within societies it is effectively fixed. While many renderings of this hypothesis exist, the following provides an excellent sense of what most have in common (1924 Biennial Report of the California State Board of Prison Directors: 9):

> The tremendous increase in the population of the prisons should not be taken as indicative of any abnormal increase of crime, but rather as illuminative of an abnormal increase in the states [sic] population. In fact, it may be accepted as fairly axiomatic that the ratio of prison population to free population in this, or any other American commonwealth, approximates the static.

Emile Durkheim's (1964a, 1964b) works are probably the best-known efforts to address this "axiomatic" relationship. According to Durkheim, criminal behavior is socially defined and to be expected, even in "a society of saints" (1964b:68–9). Crime, in turn, serves a number of im-

* Direct all correspondence to Richard A. Berk, Social Process Research Institute, University of California, Santa Barbara, CA 93106.

Thanks go to Kathy Stathopoulos and Amy Kaiser for their help in collecting and preparing the historical data used in this analysis. We also benefited enormously from the comments of John Berecochea, Stanton Wheeler, and Alfred Blumstein on an earlier draft of this paper. Funding for the research was provided by the National Institute of Justice (Grant No. 78-NI-AX-0093).

portant functions via the ritual of punishment that reinforces social integration (1964a:105–10). Finally, the level of punishment responds to competing processes that, in an equilibrating manner, contain the level of punishment within appropriate bounds; social definitions of crime change in order to have neither too few nor too many people sanctioned by "repressive" law. "What is normal, simply, is the existence of criminality, provided that it attains and does not exceed, for each social type, a certain level . . ." (1964b:66).

Durkheim's approach is particularly noteworthy because it addresses ongoing, dynamic processes. Stable levels of punishment can be derived, however, from rather different perspectives. For example, economists (e.g., Becker, 1968; Stigler, 1970; Votey and Phillips, 1973) have developed static equilibrium models in which local jurisdictions equate the marginal benefits of law enforcement to the marginal costs. In the short run, consequently, there is a stable level of punishment that maximizes aggregate well being. Yet, such models are necessarily silent on how temporal adjustments occur; they are static "snapshots" of relationships already in equilibrium.

When one grants that political jurisdictions may seek an optimal balance between enforcement and crime, the door is opened to views in which some stable level of punishment responds not to a functional requirement for society as a whole, but to the actions of a ruling elite. In other words, the equilibrating process of punishment becomes a means to an end, and that end can also be defined through "ruling class" interests. For example, in a recent historical study of crime in early modern Seville, Perry (1980) documents the trade-offs between incarcerating large numbers of the local underclasses, providing the requisite numbers of military recruits, and maintaining the necessary labor pool for Seville's economic base. Similarly, Ruggiero's (1980) study of crime in 14th century Venice shows that there were significant opportunity costs when incarceration rates (and rates of mutilation) were increased. In both instances, ruling elites recognized these costs and tried to strike a proper balance between crime and enforcement. These studies are, of course, fully consistent with more explicit formulations (e.g., Rusche and Kirchheimer, 1968; Ignatieff, 1978; Foucault, 1980) in which the kinds and/or levels of punishment found in Western society since the Middle Ages respond in rather stable ways to the prevailing economic system.

It appears, therefore, that the "axiomatic" relationship between the number of people punished and the total number of people in the population can be derived from a variety of equilibrating formulations. It is equally apparent that many scholars believe that the proportion of people punished is relatively constant over time. What is the evidence? In fact, there have been few studies in which the stability of punishment hypothesis is tackled head-on. The historical studies cited above, for example, address equilibrating tendencies only in passing and never directly consider empirical estimates of the punishment rate over time (partly because of inadequate data). Similarly, the studies within the neoclassical economic tradition have focused on static models in which an equilibrium is assumed. Perhaps the most visible efforts to address the rates of punishment over time appear in a series of articles by Alfred Blumstein and several of his colleagues (Blumstein and Cohen, 1973; Blumstein, Cohen, and Nagin, 1977; Blumstein and Moitra, 1979). Building heavily on Durkheim's views, these studies purport to show that rates of punishment are quite stable over relatively long intervals in a variety of Western countries and throughout political jurisdictions in the United States. However, these studies have been critized for the particular historical periods analyzed (e.g., Cahalan, 1979), definitions of the punishment rate used (e.g., Waller and Chan, 1974), and the statistical procedures employed (e.g., Rauma, forthcoming). We share these concerns, but believe that until a general model for punishment rates is formally derived, properly specified, and directly tested, a compelling consideration of the stability of punishment hypothesis cannot be undertaken.

With this in mind, we will develop below a dynamic model of punishment that includes "Blumstein's model" as a special case. Then, using data from California for the period 1851 (when the state prison system began) to 1970, we will formally test for stability in rates of punishment and consider the substantive implications of the resulting equations. While our approach has much in common with the important and recent work by sociologists studying the macrodynamics of crime (Land and Felson, 1976; Cohen and Felson, 1979; Cohen, Felson, and Land, 1980), we will focus on punishment and the possible role of equilibrating processes.

A DYNAMIC MODEL TO TEST THE STABILITY OF PUNISHMENT HYPOTHESIS

In "A Theory of the Stability of Punishment," Blumstein and Cohen (1973) propose the following expression for the constant relationship between the number of people punished and the number of people in the population:

$$P_t/Pop_t = K \qquad (1)$$

where P_t is the number of people punished at time t, Pop_t is the number of people in the population at time t, and K is a constant. Stated discursively, the fraction of the population punished at any given time is a constant.

In later work (particularly Blumstein, Cohen, and Nagin, 1977), Blumstein and his colleagues develop a formal model of the underlying equilibrating processes through which the constant "K" materializes. With "punishment" operationalized as the number of people incarcerated, they rely heavily on Durkheim's notion that the boundaries between criminal and noncriminal behavior are continually altered to produce a relatively constant proportion of people being punished. Three differential equations follow that link the rate of change in the number of people incarcerated, the rate of change in the number of people defined as criminal, and the rate of change in the number of people defined as law-abiding. With the aid of several simplifying assumptions (see, for example, Blumstein, Cohen, and Nagin, 1977:324), they conclude that the proportion of people incarcerated should, over time, follow a stationary, second-order autoregressive process.[1] This formal result and the necessity of allowing for random perturbations in turn imply an empirical test of the initial differential equation formulation. If the proposed equilibrating processes are correct, it should be possible to estimate a univariate, ARIMA, time-series model for the proportion of people incarcerated over time (as shown in Footnote 1) that is stationary and contains statistically significant autoregressive parameters at lags of one and two (only). Such tests are undertaken, and the required model surfaces.

There are at least three reasons for being uneasy with this strategy (cf. Greenberg, 1978:348–62). First, there is no explicit role for exogenous variables. Rather, the model depicts adjustments that follow when too few or too many individuals are punished. Why such imbalances occur cannot be addressed, and the richness implied by the traditions reviewed above cannot be incorporated; there is no way to assess the impact of political unrest, economic conditions, wars, and the like. Second, the model is seriously underidentified, and unique parameter estimates cannot be obtained without a number of additional assumptions and information from external sources. Thus, much of the model cannot be empirically tested, and many of the most important causal relationships are not subject to falsification. Third, given these and other problems, Blumstein and his colleagues attempt to verify their formulation through the derived stationary, second-order, autoregressive process. However, while a reasonable fit is appar-

[1] Such a model has the general form of $Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \epsilon_t$, where Y_t is the fraction of the population being punished at time t, the betas are essentially regression coefficients, and the epsilon is an error term meeting the usual ordinary least squares assumptions. Broadly stated, by "stationary" one means that the longitudinal process is "the same" regardless of what time intervals are examined. For an excellent introductory discussion of these and related issues see Pindyck and Rubinfeld (1981:493–524).

ently obtained, the model itself is not directly tested, only its implications. Moreover, low-order, autoregressive processes are very common in time-series data and can result from any number of underlying mechanisms unrelated to the proposed differential-equation framework. At best, the test is weak.

These and other concerns demonstrate the need for a more general and flexible model of punishment. This model should be dynamic, should include parameters for the rôle of substantively meaningful exogenous variables, and should explicitly represent any equilibrating processes. Such desiderata favor restricted versions (described below) of rational polynomial functions of the following form:

$$A(L) \ln Y_t = B(L) \ln X_t + C(L) \ln a_t \quad (2)$$

where A(L), B(L), and C(L) are polynomials in the lag operator L, Y_t is an endogenous variable, X_t is an exogenous variable, and a_t is an error term. The application of log transformations, under certain specifications, has the convenient result of allowing one to interpret Equation 2 in terms of growth rates. In any case, restricted forms of such models are finding increasing use in economics, and a general discussion of their properties can be found in Davidson et al. (1978) and Hendry and Mizon (1978).

From Equation 2 alone, it is not readily apparent how one can formulate models for the level of punishment; the requisite restrictions must be discussed. We turn now to these issues and, for purposes of exposition, will proceed with simplified versions of the models we will ultimately use.

Model I: A Fixed Target with Short-Term Deviations

Consider first a simple model with the number of people punished at time t (P_t) and the number of people in the population at time t (Pop_t) as its two generic variables. One set of restrictions on Equation 2 yields the following equation:

$$\ln(P_t/P_{t-1}) = \alpha_0 + \alpha_1 \ln (Pop_t/Pop_{t-1}) + \alpha_2 \ln (P_{t-1}/Pop_{t-1}) + \ln a_t \quad (3)$$

Equation 3 expresses the growth rate[2] of punishment as a function of the growth rate of the population, the proportion of the population being punished, and an error term. Demographic forces alone imply that the two growth rates are perhaps positively related; an increase in the growth rate for the overall population leads to an increase in the growth rate of punishment. Consequently, α_1 should be positive. The proportion of people being punished is assumed to affect the growth rate of punishment, one time period later. In Equation 1 this proportion is what Blumstein and others have assumed to be constant over time; but, at this point, we are simply asserting that as the proportion increases, the growth rate of punishment decreases (and vice versa). Therefore, α_2 should be negative. In this context, it is useful to conceptualize the proportion as the "social control ratio" and a "target" for the system. When the social control ratio is too low, the growth rate of punishment will increase. When the social control ratio is too high, the growth rate of punishment will decrease. This implies that there is pressure to keep the social control ratio constant. That is, there is a tendency toward an equilibrium ratio of punishment to the population, and Equation 3 as a whole represents an equilibrating process. Put in other terms, Equation 3 is a "control rule" indicating how the punishment growth rate is altered in response to exogenous changes in the population growth rate so that the target level of social control is maintained. And since it is the punishment growth rate that can be altered within the system, it is the punishment growth rate whose temporal variation is to be "explained." Finally, for now we will assume that the error term fulfills the usual OLS assumptions.[3]

[2] For small changes, the difference between the log of a variable at time t and the log of that variable at time t−1 very closely approximates the percentage change in that variable (Theil, 1978:188–91). Thus "change in the logs" has a growth-rate interpretation that is commonly used.

[3] The reasons for the specific lags on the two right-hand-side variables can be found in the references cited above (especially Davidson et al., 1978:679–82). In brief, the formal derivation of the model implies that one cannot employ the social

With these preliminaries behind us, the implications of Equation 3 can be considered in more detail. To begin, assume a "steady state" in which the rates of growth are constant. That is, each growth rate is fixed at some value, and that value is what may be "typical" for that variable over time in a given jurisdiction (or society). Restated, the steady state can be defined in terms of constant rates of growth (or more broadly, constant rates of change).[4]

Under the steady-state conditions, one can return to Equation 3 and insert typical (and fixed) growth rates for punishment and the population. In practice, the mean growth rate for each is often appropriate. Then, one can solve for the social control ratio and obtain

$$\exp\{-(\alpha_0/\alpha_2)-(\alpha_1/\alpha_2)\,\overline{\text{Pop}} + (1/\alpha_2)\overline{\text{P}}\} = (\text{P/Pop}) = \text{The Target (K)} \qquad (4)$$

The bars over the two variables indicate that the constant growth rates for population and punishment have been inserted, and exp denotes the base for natural logarithms. All else is defined as in Equation 3, although we no longer need the "t" subscripts (i.e., in the steady state).

In Equation 4, each coefficient and constant growth rate plays a role in determining the value of the social control constant. For example, assume that the constant growth rate for punishment is .095 (i.e., the constant ratio of P_t to P_{t-1} is 1.10), and the constant growth rate for population is .048 (i.e., the constant ratio of Pop_t to Pop_{t-1} is 1.05). Then if α_0 equals .01, α_1 equals 1.0 and α_2 equals $-.01$, the (constant) social control ratio in the steady state is .02. That is, 2% of the population is being punished. Should any of the growth rates or coefficients differ, another value for the constant would be obtained. For example, jurisdictions characterized by more rapid population growth in the steady state may require more social control. Jurisdictions with higher rates of immigration, for instance, may initiate more social control to regulate the behavior of "unsocialized" newcomers. Therefore, the steady-state, social-control ratio will be greater. The general point is that should the constant materialize in a given empirical setting, one has the tools to explain its magnitude and why it differs from constants in other settings.

Equally important, the constant shown in Equation 4 (the target) is precisely the same expression used by Blumstein and his colleagues to represent in Equation 1 the constant ratio of the number of people punished to the number of people in the population. However, we have derived the constant through the concept of a steady state coupled with Equation 3. We are basically arguing that discussions of the stability of punishment hypothesis imply that "society" is operating "as if" the control rule stated in Equation 3 were in place. Once the control rule is expressed in formal terms, one can extract new insights about the proposed equilibrating tendencies. In particular, Equation 4 suggests that it may be instructive to view "Blumstein's constant" as the constant ratio of punishment to the population in the steady state of constant growth (or more broadly, constant change).[5]

It is also important to examine the implications of the steady state for the time path of punishment. In steady state, the social control ratio is fixed at some value and therefore is incorporated within the intercept of Equation 3. Yet, since the population is changing (at a constant rate), the number of people punished is changing (at a constant rate). Conse-

control ratio at a lag of zero. It becomes apparent from inspection of the unconstrained version of Equation 3 that the same variable at the same lag (i.e., P_t) would be both sides of the equal sign. However, additional lags beyond a lag of one period can be used. As for the population growth rate, its primary and essential purpose is to control for the impact of population so that its effects are not confounded with those of the social control ratio (since the social control ratio has population in its denominator). The necessary lags follow from this role, but again, additional lags can be used.

[4] For dynamic models, this is the standard definition of a steady state. Also, in steady state, the error term is assumed to be zero.

[5] For a more formal derivation of the relationship between the steady-state constant and the appropriate dynamic model, see Davidson et al. (1978:679–81). The main insight to be gained from the derivation is the precise nature of the constraints linking the general rational polynomial form, the equilibrating model, and the steady-state constant.

quently, the time path for the growth rate is a horizontal line, as can be seen in Figure 1(a). However, a constant rate of growth implies a nonconstant, nonlinear time path for punishment. If the growth rate of punishment is positive, the time path for punishment will look something like the time path shown in Figure 1(b). One will have a smooth monotonic trend over time as the steady-state time path of punishment.

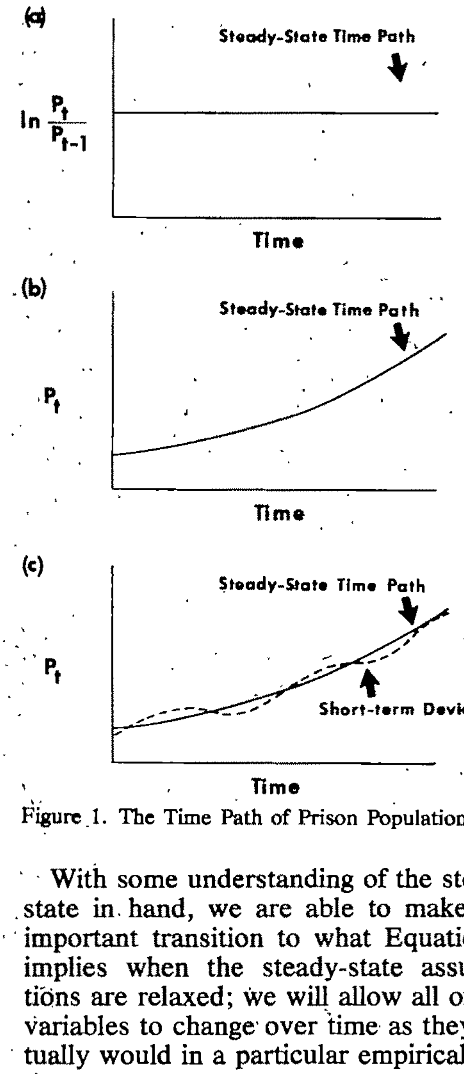


Figure 1. The Time Path of Prison Population (P_t)

With some understanding of the steady state in hand, we are able to make the important transition to what Equation 3 implies when the steady-state assumptions are relaxed; we will allow all of the variables to change over time as they actually would in a particular empirical setting.

For purposes of exposition, assume that the variables in Equation 3 have been in steady state for some time. Thus, the number of people punished has the smooth time path shown in Figure 1(b). Now, at some time point t, we relax the assumption that the population is growing at a constant rate. The result is that the change in the growth rate of punishment for that period will not represent the steady state, but deviations from that steady state. In other words, the time path for punishment will no longer be the steady-state time path. However, in the next time period, the values for the number of people punished and the number of people in the population (reflecting non-steady-state growth rates) will appear in the social control ratio. Assuming the equilibrating process is captured in the social control ratio (i.e., α_2 is negative), the growth rate of punishment now will be adjusted toward the steady-state rate of growth. If α_2 is large, the adjustment for the one-period deviation from the steady state will occur rapidly. If α_2 is small, the adjustment will require a longer interval. In short, changes in the population that do not reflect the steady-state growth rate will introduce short-term deviations from the steady-state time path of punishment much as shown in Figure 1(c), which shows short-term deviations from the steady state under the assumption that population continues to change from period to period, but not at the steady-state rate.

Model II: Model I with Shifts in the Target

Our first model rests on the idea that the social control ratio provides a fixed target for the dynamic system. The target is defined as a steady state toward which the system aims (with constant rates of growth). However, one could easily imagine situations in which there are discrete changes in the level of social control required in a jurisdiction; that is, there are discrete changes in the target. For example, in periods of war, less social control may be required (if the war is popular). During serious economic upheavals more social control may be required. This implies the need to develop more complicated control rules in which the target for the system shifts in a discrete manner. One could still have an equilibrating system, but the target level of social control is no longer rigidly fixed over time. At the very least such a formulation seems con-

sistent with the spirit of the Durkheim-Blumstein approach, although an unchanging, steady-state solution will not materialize. In other words, the notion of an equilibrating process may be basically correct despite the absence of a single, steady-state, social control constant. In short, one is perhaps able to give the Durkheim-Blumstein perspective the benefit of the doubt.

Consider again Equation 3, but with a new variable, W_t:

$$\ln(P_t/P_{t-1}) = \beta_0 + \beta_1\ln(Pop_t/Pop_{t-1}) + \beta_2\ln(P_{t-1}/Pop_{t-1}) + \beta_3W_t + \ln a_t \quad (5)$$

W_t is a binary variable coded 0 in times of peace and coded 1 in times of war. Assuming that β_3 is negative, during times of war there will be a drop in the intercept of Equation 5 that translates into a drop in the growth rate of punishment; the growth rate will decrease in a discrete fashion (although it may still be positive).

At first glance, it might seem that the binary variable for periods of war has introduced another kind of short-term deviation from the steady-state time path of punishment. However, after solving Equation 5 for the steady state as before, one obtains

$$\exp\{-(\beta_0/\beta_2)-(\beta_1/\beta_2)\overline{Pop}+(1/\beta_2)\overline{P} -(\beta_3/\beta_2)W_t\} =(P_{t-1}/Pop_{t-1}) \quad (6)$$

Equation 6 indicates that, in periods of war, the social control ratio is shifted; the target for the equilibrating system is altered. Using the earlier numerical example, if β_3 is $-.02$, the new constant in steady state is .003 (or .3%).[6] In this instance, the need for social control has been markedly reduced from punishing 20 people per thousand in the population to only 3 per thousand in the population. In summary, there is still an equilibrating process, but the steady-state social control ratio need not be unchanging.

[6] In practice, the α's from Equation 4 will differ from the β's in Equation 6. Hence, disparities between the social control constants from the two equations will not be solely a function of the impact of β_3. Also, because the war dummy variable requires a subscript, the social control ratio does as well.

Model III: Model II with Continuous Changes in the Target

If the social control target can be shifted in response to binary variables, a natural question is whether a model can be formulated in which the social control target moves "continuously" (i.e., for each period in discrete time). Consider the following:

$$\ln(P_t/P_{t-1})=\gamma_0+\gamma_1\ln(Pop_t/Pop_{t-1}) +\gamma_2\ln(P_{t-1}/Pop_{t-1})+\gamma_3W_t +\gamma_4\ln C_t+\ln a_t \quad (7)$$

Equation 7 is the same as Equation 5 except that a new variable has been added. That new variable, C_t, is a continuous variable representing the capacity of the jurisdiction's prison system. Presumably, this capacity can change over time as a function of new prisons, additions to existing prisons, and/or crowding more prisoners into the existing facilities (and hence, redefining capacity). From Equations 7, one can see that, if γ_4 is positive, an increase in prison capacity will translate into an increase in the growth rate of punishment.

The implications for the steady-state social control target can be found by once again solving for the social control ratio, assuming constant rates of growth. Thus;

$$\exp\{-(\gamma_0/\gamma_2)-(\gamma_1/\gamma_2)\overline{Pop}+(1/\gamma_2)\overline{P} - (\gamma_3/\gamma_2)W_t-(\gamma_4/\gamma_2)C_t\}=(P_{t-1}/Pop_{t-1}) \quad (8)$$

Since the impact of prison capacity in Equation 7 has not been formulated in growth-rate terms, one cannot speak about its steady-state constant rate of growth. In other words, the steady state has absolutely no implications for the permissible values of C_t. One can see from Equation 8 that, as a consequence, the social control target will be "continuously" shifted as a function of prison capacity, even in steady state. It is important to stress that there is still in principle a target for the system. However, the target will vary in each time period (according to variation in prison capacity). We have therefore allowed for a still broader generalization that may be consistent with the spirit of the Durkheim-Blumstein approach. There remains an

equilibrating process in the absence of a single, steady-state social control ratio. Again, we are giving the stability of punishment hypothesis the benefit of the doubt. There may be no apparent stability of punishment, but equilibrating forces may well be at work.

To briefly summarize, there are an enormous number of equilibrating models within a dynamic-difference equation framework (Chiang, 1974: 549–621). By capitalizing on the particular properties of restricted, rational, polynomial functions, we are able through Model I to characterize one kind of equilibrating process that is consistent with the Durkheim-Blumstein perspective, and which in steady state yields the traditional stability of punishment hypothesis. However, the Durkheim-Blumstein equilibrating process can be generalized as in Model II to allow for a set of discrete steady-state solutions; punishment may stabilize at different levels depending on different historical circumstances. Finally, the Durkheim-Blumstein equilibrating process can be generalized still further, as in Model III, to allow for continually changing steady-state solutions; punishment levels will never stabilize despite ongoing equilibrating adjustments. Thus, within the Durkheim-Blumstein approach there are actually two related issues: is there evidence of equilibration, and in what sense are punishment levels stable?

DATA, VARIABLE CONSTRUCTION, AND STATISTICAL PROCEDURES

In principle, the Durkheim-Blumstein perspective could be tested in a wide variety of jurisdictions, societies, and historical periods. However, the moment that one moves beyond the very simplest kinds of dynamic formulations, a number of practical considerations intervene. First, specifications consistent with the dynamic models described above may well require more than the two variables of punishment and population. Therefore, a setting (or settings) for the empirical work must be found in which a rich array of variables is easily obtained. Second, for powerful tests of the relevant hypotheses, a sufficient number of longitudinal observations must be available. In the case of yearly data, for example, this is a significant constraint because one may need well over 50 years of historical material. Finally, an accurate specification of the appropriate dynamic equation requires more than the broad theoretical formulation described above. One needs, in addition, considerable grounded knowledge of the setting from which the data come and some historical understanding about important trends and events. For these and other reasons, we settled on the State of California and the historical period from the opening of its prison system in 1851 to 1970.

Perhaps the most important and thorny problem with respect to the models being proposed is how to define, operationalize, and measure punishment (Melossi, 1980). We will begin by using the total number of people incarcerated in state prisons in California as a measure of punishment, and the number of people incarcerated divided by the total state population as our indicator of social control. This is basically the way in which Blumstein and his colleagues chose to address the stability of punishment hypothesis, and thus provides a good place to start. Alternative measures will be considered later.

Turning to our statistical procedures, we will draw heavily on the works of Davidson et al. (1978) and Hendry and Mizon (1978). These authors propose direct tests for restricted versus unrestricted versions of rational, polynomial functions such as those we have been discussing. In brief, one first builds a constrained version of the model with a sufficient number of lagged regressors to produce white noise residuals from ordinary least-squares procedures (which basically means that all systematic variation is accounted for). In our case, the constrained specification will include the social control ratio along with several other terms. A simple variant of this type of model was discussed earlier as Equation 3. One then relaxes the restrictions of interest, which leads to specifications much like Equation 2: ratios of variables are "broken up." For us, this means relaxing the restrictions associated with the equilibration process. Finally, an F-test for the difference be-

tween the error sum of squares for the constrained and unconstrained equations can be used to determine if the constrained specification should be rejected. A statistically significant F-ratio indicates that the constrained formulation is misspecified and that, in particular, the equilibrating model is incorrect. However, if one discards the equilibrating process, one also discards the steady-state target. One does not have to worry, therefore, about whether the social control target is fully fixed, shifted in a discrete manner, or constantly changing, and what these variants mean for the stability of punishment hypothesis. Only if one fails to reject the equilibrating process need one be concerned with the nature of the steady-state target. In short, the F-test provides a direct test that is in theory unambiguous; a rejection implies that there is no equilibrating process and therefore no steady-state target, be it stable or otherwise. Readers seeking a more thorough discussion of these procedures and their justifications should consult the references cited above, particularly Davidson et al. (1978:680–1).

SOME TESTS OF THE STABILITY OF PUNISHMENT HYPOTHESIS

In the analyses to follow we will begin with relatively simple specifications consistent with those considered earlier under Model I: a fixed target with short-term deviations. Recall that Model I provides for a direct test not only of the equilibrating process, but also of whether the steady-state social control ratio is really unchanging. Consequently, tests based on Model I address the traditional stability of punishment hypothesis, not the more flexible formulations (i.e., Models II and III) developed in the spirit of the Durkheim-Blumstein perspective.

The historical period of interest will begin in 1853 and end in 1970. The first two years for which data are available are excluded because there were so few prison admissions and, consequently, the growth rates were clear outliers. In addition, while data on many variables were available through 1980, some series were available only through 1970. Data on the number of individuals punished was obtained primarily from California prison reports. The incarceration series reflects the number of persons in California state prisons, including prison camps, on one day during each year, 1853–1970. The day of count was usually June 30th for the years 1853–1944 and December 31st for the years 1945–1970. The main data source for pre-1945 figures is the annual and biennial reports of the California State Board of Prison Directors. For figures beginning in 1945 see, in the main, *California Prisoners* and *Summary Statistics: Civil Commitment Program for Narcotic Addicts*, both published by the California Department of Corrections. Other unpublished data were supplied by the Department of Corrections.

"Persons in prison" during these years were primarily, but not only, persons convicted under California law and sentenced to state prison as "felons." Additionally, a few persons convicted under federal law have been housed in California state prisons since 1880. Further, since at least the 1940s, California state prisons have housed small numbers of tuberculars convicted of the misdemeanor of violating quarantine regulations ("recalcitrant tuberculars"), persons convicted of a felony in other states, persons committed to mental health authorities, and persons awaiting disposition of charges who were housed in state prisons for "safekeeping." Since the 1940s California state prisons have also housed three larger categories of persons not sentenced as "felons:" persons under civil commitment to the program for narcotic addicts operated by the Department of Corrections, a portion of those committed to the California Youth Authority as "juveniles" or "youths," and those committed to the Department of Corrections prior to sentencing for a 90-day "diagnostic period."

Population figures were gathered from the census with exponential interpolations to fill in the gaps between census years; the growth-rate formulation implies an exponential time path for population.

Recall that Blumstein and his colleagues defined punishment in terms of the number of people incarcerated in state prison. For our California data over the

period of interest, the number of people in prison varied from a low of 228 to a high of 28,531, with a mean of 6,477. Over-time patterns show (see Figure 3 below) an increasing number of people incarcerated, with the rate of increase clearly increasing as well. Consistent with this trend, the mean growth rate for the state's prison population is .04; the prison population increased at an average rate of about 4% a year. Yet, there was also considerable variation in year-to-year growth rates. Indeed, in some years the growth rate was negative (the most negative growth rate was −.18).

The total population figures for California follow a roughly similar temporal pattern, although the raw numbers are, of course, far larger. This leads to a social control ratio ranging from a low of about 70 per 100,000 people to a high of 220 per 100,000 people, with a mean of 140 per 100,000.

Figure 2 shows the social control ratio plotted over time, with punishment defined as the number of people incarcerated per 100,000 in the population (the definition used by Blumstein). Even ignoring the first few years in the series, one might be tempted to conclude that the ratio is not constant. There appears to be an overall downward drift, with dramatic valleys in the 1920s and 1940s. However, our earlier discussion should make clear that whether or not the social control ratio is stable cannot be determined by inspection. Under Model I, the steady-state ratio may well be fixed, but the observed ratio is still subject to short-term perturbations. Equally important, under Models II and III one can allow for equilibration coupled with systematic variation in the steady-state target.

Table 1 reports a formal test of Model I. The constrained specification shown at the top of the table is consistent with Model I (i.e., a fixed target with short-term deviations) with one minor modification. A single lagged value of the prison population growth rate was required to produce white noise residuals. The unconstrained specification is shown in the middle of the table and reflects a relaxation of the restrictions implied by the

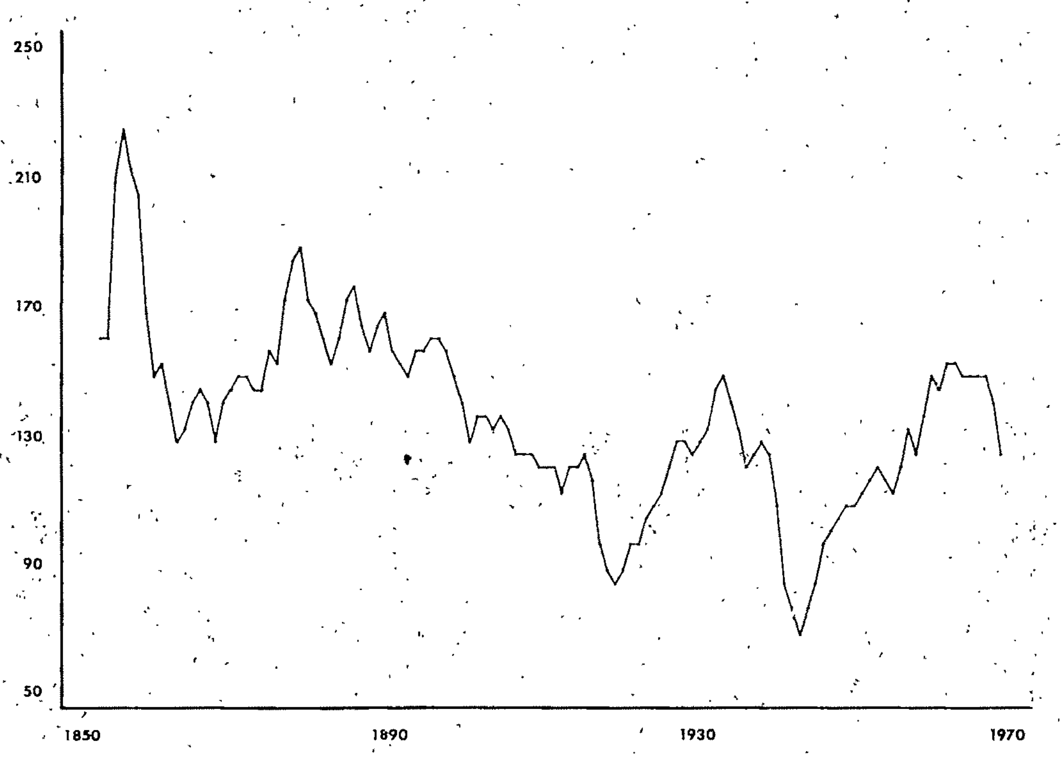


Figure 2. Prison Population in California, per 100,000, 1853–1970

Table 1. A Test of Model I for Total Prison Population

Constrained Model:
$\ln (P_t/P_{t-1}) = \alpha_0 + \alpha_1 \ln (Pop_t/Pop_{t-1}) + \alpha_2 \ln (P_{t-1}/Pop_{t-1}) + \alpha_3 \ln (P_{t-1}/P_{t-2}) + \ln a_t$

Parameter	Coefficient	t-Value
α_0	−.57	−3.02
α_1	.67	2.68
α_2	−.08	−3.02
α_3	.47	5.50
ESS = 0.428	DF = 112	R^2 = .34

Unconstrained Model:
$\ln (P_t/P_{t-1}) = \beta_0 + \beta_1 \ln Pop_t + \beta_2 \ln Pop_{t-1} + \beta_3 \ln P_{t-1} + \beta_4 \ln P_{t-2} + \ln a_t$

Parameter	Coefficient	t-Value
β_0	−.61	−3.24
β_1	.44	2.03
β_2	−.33	−1.54
β_3	.38	4.51
β_4	−.50	−5.93
ESS = 0.414	DF = 111	R^2 = .36

F test = 3.52 (df = 1,112) (for the reduction in ESS from the constrained to the unconstrained model)

equilibrating process represented by Model I. All of the regressor growth rates and the social control term have been "broken up." Recall that it was precisely these ratios that were implicated in the equilibrating process. Note also that the two specifications differ by only one degree of freedom. (This is a general result for the models we will be considering.)

The F-test reported at the bottom of Table 1 indicates that we fall just short of rejecting the constrained specification. The F-ratio is 3.52 (df = 1, 112) and we need an F-ratio of 3.92 in order to reject the constrained version at the .05 level. The probability value associated with the estimated F-ratio is approximately .07. Formally, at least, this means that we have to take a close look at the results from the constrained formulation.

Focusing on α_2 for the equilibrating term, the coefficient is statistically significant at the .05 level, and the sign is negative (as predicted). At first glance, therefore, there is some evidence for the hypothesized equilibrating process. However, in this instance, appearances may well be deceiving. To begin, the estimated correlation between the intercept and the regression coefficient for the social control ratio is .99 with the relevant condition index well over 60 (Belsley et al., 1980); there are clear warnings of very serious collinearity. According to Davidson et al. (1978:683), such collinearity between the constant and the equilibrating term is quite common in these formulations, and they demonstrate that the proper solution is to drop the constant from the model. In effect, this constrains the intercept to zero, which in a growth-rate metric is not likely to be a significant distortion. Readers familiar with Box-Jenkins (Box and Jenkins, 1976) approaches to these kinds of data will recognize that we are merely saying that the trend parameter is not required, and such readers will also recall that the absence of a trend parameter in differenced models is hardly unusual.

If one drops the constant from the model, the apparent impact of the social control ratio disappears completely and the t-value falls below 1.0 in absolute value. This is just the sort of instability one expects when one's data are highly collinear. In short, even if one chooses not to reject the constrained specification through the F-test, we can find little evidence for an equilibrating process when punishment is defined in terms of prison population. This implies a rejection of the stability of punishment hypothesis as well. However, the credibility of this conclusion depends on the credibility of our Model I formulation. We will see later that Model I is almost certainly misspecified.

One might alternatively broaden the definition of punishment to include the number of people in prison plus the number of people on parole. Surely parole is a form of punishment, and perhaps punishment rates are stable if the punishment "net" is widened. In fact, we explored this possibility with a second equation. The second series counts the number of persons in California state prisons, including prison camps, on one day during each year 1853–1970 plus the number of persons under active parole supervision of state adult parole authorities, usually on the same day during each of those years. The parole population includes both persons under commitment as "felons" and those under civil commitment as "narcotic

addicts." The latter are officially said to be on "outpatient status" rather than "parole."

Substantially all persons on outpatient status were being supervised in California. In different years, different proportions of those on parole were being supervised in another state, or had been placed in the custody of immigration authorities or of other law enforcement authorities. We have left them in the parole counts because they remained under the "constructive" jurisdiction of California adult parole authorities.

One further anomaly is worth noting. Those persons committed to the California Youth Authority housed in state prisons were counted as part of the prison population. Upon parole, such persons are supervised by Department of Youth Authority parole agents, not those of the Department of Corrections. For this reason, and because we could not reliably distinguish them from Youth Authority parolees who had not been housed in state prisons, we did not include them in the parole populations.

Again, the F-test failed to reject the constrained model, and again the coefficient for the equilibrating term with the collinear intercept deleted was not statistically significant ($t = .54$). In the interest of saving space we do not provide a summary table.[7]

One can also define punishment in terms of the number of people admitted to prison. Focusing first on admissions for felony offenses committed while not under supervision of the State Corrections System (i.e., on parole), one has the advantage of working with a cleaner definition of punishment; prison admissions are not confounded with the length of sentences being served. In any case, the mean growth rate of .03 is nearly the same as the earlier growth rates for persons on-hand.[8]

Table 2 provides the results for the constrained and unconstrained versions of Model I applied to the growth rate of "new" prison admissions. The constrained specification no longer requires a lagged value of the endogenous variable to achieve white noise residuals, and there appears to be some role for the social control term. However, this is once again a function of the very high collinearity with the intercept ($r = .99$, condition index > 60); and, more importantly, the F-ratio of 12.24 ($df = 1,114$) clearly rejects the constrained formulation at the .05 level. In short, we can find no evidence for equilibration when punishment is defined in terms of the number of people admitted as "new" offenders.

Table 2. A Test of Model I for "New" Prison Admissions

Constrained Model:
$\ln(A_t/A_{t-1}) = \alpha_0 + \alpha_1 \ln(Pop_t/Pop_{t-1}) + \alpha_2 \ln(A_{t-1}/Pop_{t-1}) + \ln a_t$

Parameter	Coefficient	t-Value
α_0	−1.04	−2.98
α_1	1.81	2.92
α_2	−0.13	−2.97
ESS = 2.66	DF = 114	$R^2 = .09$

Unconstrained Model:
$\ln(A_t/A_{t-1}) = \beta_0 + \beta_1 \ln Pop_t + \beta_2 \ln Pop_{t-1} + \beta_3 \ln A_{t-1} + \ln a_t$

Parameter	Coefficient	t-Value
β_0	−1.45	−4.38
β_1	1.58	3.74
β_2	−1.34	−3.24
β_3	−0.31	−4.73
ESS = 2.40	DF = 113	$R^2 = .18$

F test = 12.24 (df = 1,114) (for the reduction in ESS from the constrained to the unconstrained model)

[7] One might consider still broader definitions of punishment, including such things as the number of people on probation or even the number of people in mental hospitals. However, such data were not available over the full period of interest.

[8] The series records all persons received by California state prisons with "new" felony sentences from California courts each year, 1853–1970. A "new" sentence is one received when the person is not at the time serving a felony sentence under the jurisdiction of a state prison or adult parole authorities. Such admissions are only a portion of prison receptions, as implied by the earlier discussion of persons in California state prisons. Omitted are persons received with other than felony sentences (e.g., Youth Authority wards, narcotic addicts under civil commitment, "diagnostic" cases, "recalcitrant tuberculars," federal and other states' felons, mental patients, "safekeepers") and persons already serving a sentence under the jurisdiction of state prison and adult parole authorities who are recommitted with or without a new conviction (e.g., returned parolees and outpatients and returned escapees). Also omitted are prisoners returned from mental hospitals and prisoners returned from court.

Suppose now that a broader definition of admissions is used: new felon admissions plus the other major types of admissions received by the California state prisons each year, 1853–1970. The latter include returned escapees, many of whom received an additional felony sentence; narcotic addicts under civil commitment, including both "new" commitments and persons returned from outpatient status with and without new convictions; returned parolees with and without new convictions; and Youth Authority wards, including both "new" commitments and those returned as violators cases. Prisoners returned from mental hospitals and courts have been omitted.[9] Despite the use of this broader definition of punishment, we still find no evidence of equilibration. While the F-ratio is not statistically significant, the t-value for the social control ratio is well under 1.0 in absolute value once the highly collinear constant is removed.

In summary, the four initial equations based on Model I provide virtually no support for any equilibrating process and, therefore, virtually no support for the stability of punishment hypothesis. In one instance there was a clear rejection through the F-test, and in the other three the social control ratio by itself proved unimportant. Yet there is some reason for being dissatisfied. Perhaps most significantly, the F-tests and t-tests did not produce fully consistent signals about the equilibrating process being examined. At the very least, this implies the need for greater statistical power and the real possibility of specification errors. In addition, even if Model I is in fact a total failure, there remains the prospect that alternative equilibrating models are appropriate in which the steady-state target is shifted in a discrete or continuous fashion. Model II and Model III allow for such complications and may also be consistent with the spirit of the Durkheim-Blumstein perspective.

After reviewing a number of historical accounts of the California prison system (e.g., Lamott, 1961; Miller, 1980; Yaley, 1980), primary source materials (e.g., reports of the California Board of Prison Directors, 1852–1930), and analyses of the development of prison systems in other settings (e.g., Ignatieff, 1978; Foucault, 1980), a combination of Model II and Model III seemed most promising. On a priori grounds, therefore, the following new exogenous variables were constructed:

1. the proportion of the population living in urban areas;
2. the number of people in the military for the U. S. as a whole;
3. the number of prisons in operation;
4. a binary variable for the "invention" of probation;
5. a binary variable for the passage of the Probation Subsidy Act;
6. a binary variable for the "invention" of parole;
7. a binary variable for the change to indeterminate sentencing;
8. a binary variable for the change to "split-sentencing;"
9. a binary variable for the advent of the California Youth Authority;
10. a binary variable for the introduction of the County Diagnostic Program;
11. a binary variable for the opening of the California Rehabilitation Center for addicts;
12. a binary variable to capture five depressions that occurred over the period of interest;
13. a binary variable to represent the period when the New Deal programs of the 1930s were in place.[10]

[9] Including persons both as "new" admissions and as readmissions obviously results in counting many persons more than once. (Even the series of "new" felony admissions counts some persons more than once, of course, but typically with a break in the jurisdiction of state prison and adult parole authorities.) This seems justified because each admission or readmission represents another exercise of incarceration authority.

[10] For readers interested in the historical details, we provide the following elaborations. First, "new" prisons include San Quentin (1852), Folson (1880), California Institute for Women (1932), California Institute for Men-East (1941), Correctional Training Facility–South (1946), Correctional Training Facility–Central (1951), Duel Vocational Institution (1953), California Medical Facility (1955), Correctional Training Facility–North (1958), California Mens' Colony (1961), California Correctional Center (1963), and the Sierra Conservation Center (1965). Second, while parole was officially adopted in 1893, we used 1907 as the starting data (subject to empirical checks described below) because that was when the use of parole became common. Third, the begin-

The number of people living in urban areas was included because of the common observation that urbanization is associated with higher crime rates and with increases in the number of people subject to sanctions from the criminal justice apparatus. We therefore expected to find a positive impact on the growth rate of punishment, whatever its definition. The number of people in the military was included because military personnel, by and large, are removed from the population at risk. This is especially important given the fact that it is primarily young men who serve. We anticipated a negative relationship with punishment.

The role of the number of prisons was less clear. In particular, the number of prisons is a rather imperfect measure of the real effect we hoped to capture: prison capacity. Prison capacity is an extremely slippery concept; capacity can be altered easily by crowding more prisoners into existing facilities. Moreover, better measures (e.g., the number of square feet in state prisons) were not available over the entire historical period. Nevertheless, we expected to find a positive relationship with the punishment growth rate.

All the binary variables were coded 1 when the variable was "switched on" and 0 when the variable was switched off. Eight of these binary variables represent conscious changes in public policy that might affect the growth rate of punishment. Yet it is important to stress that each alteration no doubt reverberated throughout the criminal justice apparatus in a very complex fashion. For example, it was not at all clear how the change to indeterminate sentencing would alter the rates of punishment. In short, with a few exceptions, we had no a priori expectations about the signs of any policy-related causal effects.

There is also the problem of which policy changes to consider. In principle, the possibilities are almost endless, since, in addition to direct efforts to alter prison practices, virtually any change in almost any part of the criminal justice apparatus could affect rates of punishment. We have tried to include policies that on a priori grounds were plausible candidates, but our set is certainly not exhaustive.

The binary variable for the five depressions was meant, in part, to capture a decreased demand for labor, since it has been widely argued that rates of punishment will be higher in such periods (e.g., Rusche and Kirchheimer, 1968). Perhaps a better measure would have been unemployment figures, but such data were not available before 1890.

Finally, the binary variable for New Deal programs was included to reflect substantial government efforts to provide public employment and income supports. We anticipated that the rates of punishment would decline if public employment increased the demand for labor. One might also argue that under certain circumstances income supports reduce crime rates and hence rates of punishment (e.g., Berk et al., 1980a).

It is important to stress that, even if the exogenous variables we were able to measure lead to specifications that are funda-

ning date for indeterminate sentencing was 1917, although we used 1918 as the date (subject to empirical checks) when virtually all new prisoners were sentenced in this manner. Fourth, probation was introduced in 1903, but only used extensively after 1906. Thus, 1906 became our start date (subject to empirical checks). Fifth, the Probation Subsidy Act, passed in 1965, provided direct subsidies to counties employing probation as an alternative to incarceration, and 1965 was used as the start date (subject to empirical checks). Sixth, split-sentencing, adopted in 1928, allowed courts to sentence felons to jail as a condition of probation. 1928 was used as the start date (subject to empirical checks). Seventh, the California Youth Authority began functioning in 1941, although it did little business until 1945. We used 1945 as the start date (subject to empirical checks). Eighth, the County Diagnostic Program began in 1957 and provided for psychiatric assessment of offenders after conviction but before sentencing. 1957 was used as the start date (subject to empirical checks). Ninth, the California Rehabilitation Center began receiving civilly committed heroin addicts in 1961. We used 1961 as the start date (subject to empirical checks) and counted the Center as a new policy rather than a new facility. Tenth, figures for the number of Californians in the military were not available, hence the use of the United States figures. The assumption, therefore, is that the California and U.S. figures are highly correlated. The fact that the numbers differ substantially in magnitude will then only affect the intercept. Eleventh, depressions in California were defined (where possible) as years with double digit unemployment in the United States, included 1855–1856, 1873–1877, 1893–1898, 1921, and 1931–1940. Finally, New Deal programs were deemed in effect between 1935 and 1940 (subject to empirical checks).

mentally sound, a large number of intervening processes have been neglected. For example, we have no measure of crime rates and no measures of conviction rates. (Neither was available for the early years of the time series.) Undoubtedly, crime and conviction rates respond to at least some of the regressors discussed above, and, in turn, influence rates of punishment. However, our neglect of intervening mechanisms is benign so long as one recognizes that our equations may still be fully valid as reduced forms. In other words, our specifications will not lead to biased estimates of causal effects if intervening variables are ignored.

With the discussion of our new regressors behind us, there remains the question of precisely what functional form and lags should be used when the equations are estimated. Here we proceeded more inductively. We examined the cross-correlations between the residuals of the models estimated above and the prewhitened (where appropriate) transformations of the potential regressors. This strategy is fully consistent with the specification strategies recommended by Box and Jenkins (1976: Chapter 11). In essence, we were searching for earlier specification errors that might be remedied by including our new regressors in the proper functional form with the proper time lags. For example, while the log of the number of prisons showed little promise, the growth rate in the number of prisons seemed quite viable. In the end, we settled on specifications based on both Model II and Model III; we allowed for discrete and "continuous" changes in the steady-state target.

How well does the proposed equilibrating process fare in the new equations? Table 3 presents the results for an equation in which punishment is defined as the total number of people in prison. To begin, it is apparent that, compared to Model I, we are fitting the data far better; Model I was a misspecification. The residuals for both equations in Table 3 are white noise, and the R_2 in the constrained version, for example, has increased from .34 to .69 (a statistically significant disparity). In the absence of the highly collinear intercept, there is also a hint of an effect for the social control ratio. However, when one constructs the unconstrained model, the F-test provides clear evidence for discarding the constrained specification ($F = 5.67$, $df = 1,97$). We once again reject the hypothesis of an equilibrating process. Note, however, that had we failed to reject the equilibrating process, the traditional stability of punishment hypothesis would not have been supported; only Model I had the potential of supporting the traditional hypothesis.

When the number of people on parole is added to the number of people in prison, the results are almost identical. We manage to fit the data better than we did under Model I, and the F-ratio of 4.49 rejects the constrained specification at the .05 level. In the interest of space, no table is provided.

The story is much the same when punishment is defined in terms of the number of "new" admissions to state prisons. Despite a better fit compared to Model I, an F-ratio of 15.27 leads to an unambiguous rejection of the constrained model at the .05 level. When the broader definition of admissions is used, the same conclusions follow. An F-ratio of 7.48 provides for a clear rejection of the constrained equation. Again, no tables are provided.

In summary, it appears that we are better able to fit the data when we employ specifications based on Models II and III rather than Model I. To confirm this observation we undertook a number of F-tests, contrasting, for a given punishment time series, the simple (i.e., Model I) and more complicated formulations (i.e., Models II and III); indeed Model I is but a constrained version of Models II and III. In each case, we rejected Model I as a misspecification. It is also apparent that, under specifications consistent with Models II and III, we definitively reject the constrained equilibrating formulations. The F-ratios are well in excess of values required to attain statistical significance at the .05 level. Putting these two conclusions together, the case for equilibration is further weakened. Indeed, the more convincing the specification, the more dramatic the refutation.

Table 3. A Test of Model III for Total Prison Population

Constrained Model:

$$\ln (P_t/P_{t-1}) = \alpha_1 \ln (P_{t-1}/P_{t-2}) + \alpha_2 \ln (Pop_t/Pop_{t-1}) + \alpha_3 \ln (P_{t-1}/Pop_{t-1}) + \alpha_4 \ln (U_t/Pop_t) + \alpha_5 \ln (M_t/M_{t-1}) + \alpha_6 \ln (M_{t-1}/M_{t-2}) + \alpha_7 \ln (Pris_t/Pris_{t-1}) + \alpha_8 (1-B)PR_t + \alpha_9 (1-B)PSA_t + \alpha_{10} (1-B)Pa_t + \alpha_{11} (1-B) ISL_t + \alpha_{12} (1-B)SS_t + \alpha_{13} (1-B)CYA_t + \alpha_{14} (1-B)CDP_t + \alpha_{15} (1-B)CRC_t + \alpha_{16} (1-B)D_{t-1} + \alpha_{17} (1-B)D_{t-2} + \alpha_{18} (1-B)ND_t + \ln a_t$$

Parameter	Coefficient	t-Value
α_1 (Growth rate of prison population, t−1)	.32	3.90
α_2 (Growth rate of California population)	.27	0.93
α_3 (Social control ratio, t−1)	−.002	−1.58
α_4 (Urban proportion of California population)	.003	.023
α_5 (Growth rate of the military, t)	−.02	−1.95
α_6 (Growth rate of the military, t−1)	−.02	−2.25
α_7 (Growth rate of prisons)	−.02	−0.36
α_8 (Probation)	−.05	−1.06
α_9 (Probation subsidy act)	−.03	−0.54
α_{10} (Parole)	−.05	−0.91
α_{11} (Indeterminate sentencing)	−.09	−1.57
α_{12} (Split-sentencing)	.06	1.13
α_{13} (California Youth Authority)	.15	2.81
α_{14} (County diagnostic program)	.05	0.98
α_{15} (California Rehabilitation Center)	.04	0.80
α_{16} (Depression, t−1)	.04	2.30
α_{17} (Depression, t−2)	.03	1.73
α_{18} (New Deal)	.002	0.05

ESS = .262 DF = 97 $R^2 = .69$

Unconstrained Model:

$$\ln(P_t/P_{t-1}) = \beta_1 \ln P_{t-1} + \beta_2 \ln P_{t-2} + \beta_3 \ln Pop_t + \beta_4 \ln Pop_{t-1} + \beta_5 \ln (U_t/Pop_t) + \beta_6 \ln (M_t/M_{t-1}) + \beta_7 \ln (M_{t-1}/M_{t-2}) + \beta_8 \ln (Pris_t/Pris_{t-1}) + \beta_9 (1-B)Pr_t + \beta_{10}(1-B)PSA_t + \beta_{11}(1-B)Pa_t + \beta_{12}(1-B)ISL_t + \beta_{13}(1-B)SS_t + \beta_{14}(1-B)CYA_t + \beta_{15}(1-B)CDP_t + \beta_{16}(1-B)CRC_t + \beta_{17}(1-B)D_{t-1} + \beta_{18}(1-B)D_{t-2} + \beta_{19}(1-B)ND_t + \ln a_t$$

Parameter	Coefficient	t-Value
β_1 (Prison population, t−1)	.31	3.71
β_2 (Prison population, t−2)	−.36	−4.24
β_3 (California population, t)	.44	1.36
β_4 (California population, t−1)	−.41	−1.27
β_5 (Urban proportion of California population)	.05	2.04
β_6 (Growth rate of the military, t)	−.01	−1.42
β_7 (Growth rate of the military, t−1)	−.02	−1.90
β_8 (Growth rate of prisons)	−.008	−0.13
β_9 (Probation)	−.06	−1.26
β_{10} (Probation subsidy act)	−.001	−0.03
β_{11} (Parole)	−.05	−1.08
β_{12} (Indeterminate sentencing)	−.11	−1.97
β_{13} (Split-sentencing)	.05	0.98
β_{14} (California Youth Authority)	.13	2.50
β_{15} (County diagnostic program)	.06	1.19
β_{16} (California Rehabilitation Center)	.06	1.11
β_{17} (Depression, t−1)	.04	2.27
β_{18} (Depression, t−2)	.03	1.86
β_{19} (New Deal)	.01	0.32

ESS = .247 DF = 96 $R^2 = .71$

F test = 5.67 (df = 1,97) (for the reduction in the ESS from the constrained to the unconstrained model)

TOWARD AN EXPLANATORY MODEL OF PUNISHMENT

The null findings for equilibration do not mean that punishment growth rates vary in a totally unsystematic manner. For each of our constrained specifications we were able to account for nontrivial amounts of variance, and there were hints that certain variables were behaving as

expected. Yet it was also clear that the equations were unnecessarily complex. The equilibrating process proved unimportant and the alterations in criminal justice policy by and large performed poorly. With this information in hand, we first dropped from the constrained formulation the social control term and all the criminal-justice-policy variables with the exception of the dummy variable for the California Youth Authority. The latter was the only policy variable that proved statistically significant in each of the earlier Model II and III equations. We then proceeded with ordinary least squares and examined the cross-correlations between the residuals of each equation and the prewhitened (where appropriate) exogenous variables (both those initially excluded and those initially included). We also used the residual autocorrelation function to determine whether there was any substantial evidence of serial correlation. Respecification followed, with an eye toward developing "lean" equations that would minimize potential multicollinarity and maximize parsimony (see especially Jenkins, 1979). By and large our first specifications fared well, but in a few instances a new regressor, an additional lag of an existing regressor, or a lagged value of the endogenous variable was added.[11] Finally, we ended up with specifications that easily passed all of the conventional time-series diagnostics (Box and Jenkins, 1976:392–5).

Table 4 shows the results when punishment is defined in terms of the total number of people in prison. With eight regressors we are able to account for 65% of the variance in the prison population growth rate. For the number of people in prison, 99% of the variance is explained. However, this figure is in some sense uninteresting, since the latter almost-perfect fit depends heavily on the growth-rate formulation, where the number of people in prison is lagged by one year and implicitly brought across the equal sign with its regression coefficient constrained to equal 1.0.

Probably the most productive place to focus initially is on the lagged effect of the prison population growth rate. In the log-log form, the regression coefficient of .33 ($t = 4.07$) can be interpreted as an elasticity. Thus, a 1% change in the growth rate in one year leads to a .33% change in the growth rate the following year. The primary importance of the lagged effect of the prison growth rate is to indicate that the growth rate does not respond immediately to the impact of exogenous forces; there is some "drag" in the adjustment process. Given the inertia inherent in incarceration (e.g., sentences of longer than one year), a delay in the full response to yearly changes in external forces is to be expected. On the other hand, since the regression coefficient is well below 1.0, the adjustment is completed for all practical purposes within a relatively short period (i.e., about three years).

Consistent with our expectations, the growth rate in the state's population affects the growth rate in the prison population ($t = 3.60$). With each 1% change in the former, there is an immediate .49% change in the latter. However, with the lagged effect of the prison population growth rate, the asymptotic impact of the growth in state's population is over .70% (Pyndick and Rubinfeld, 1981:230–45). In other words, each 1% change in the regressor takes a few years to be fully felt, and the total impact roughly approximates an elasticity of .75.

The growth rate in the number of people in the military for the United States as a whole has a simultaneous effect and an effect lagged by one year (with values of −2.31 and −2.50, respectively). Both

[11] For example, in some initial efforts there was evidence in the autocorrelation function for the residuals that one or two moving average parameters were required at very short lags (e.g., at a lag of one and a lag of two years). However, since a finite moving average process can be expressed as an infinite autoregressive process (Box and Jenkins, 1976: 51–3), it was possible in practice to eliminate the serial correlation in the residuals with a small number of lagged values of the endogenous variable. That is, the theoretical, infinite autoregressive process in the residuals could be effectively represented with a finite (and low-order) autoregressive process solely in Y_t. This had the advantage of transforming a substantively uninformative noise model parameterization into a substantively useful causal parameterization. It also meant that the models could be estimated with ordinary least squares to yield consistent estimates of causal effects.

Table 4. Estimated Explanatory Model for Total Prison Population

$$\ln(P_t/P_{t-1}) = \alpha_1 \ln(P_{t-1}/P_{t-2}) + \alpha_2 \ln(Pop_t/Pop_{t-1}) + \alpha_3 \ln(M_t/M_{t-1}) + \alpha_4 \ln(M_{t-1}/M_{t-2}) + \alpha_5 \ln(Pris_t/Pris_{t-1}) + \alpha_6(1-B)CYA_t + \alpha_7(1-B)D_{t-1} + \alpha_8(1-B)D_{t-2} + \alpha_9(1-B)ND_t + \ln a_t$$

Where:

P_t = Total prison population
Pop_t = California population
M_t = Number of people in the military
$Pris_t$ = Number of prisons
CYA_t = California Youth Authority dummy
D_t = Depression dummy
ND_t = New Deal dummy
a_t = White noise error

Parameter	Coefficient	t-Value
α_1 (Growth rate of prison population, t−1)	.33	4.07
α_2 (Growth rate of California population)	.49	3.60
α_3 (Growth rate of the military, t)	−.02	−2.31
α_4 (Growth rate of the military, t−1)	−.03	−2.50
α_5 (Growth rate of prisons)	−.00+	−0.03
α_6 (California Youth Authority, t)	.16	2.97
α_7 (Depression, t−1)	.04	2.29
α_8 (Depression, t−2)	.03	1.76
α_9 (New Deal)	.00+	0.12
R^2 [for $\ln A_t$] = .99	R^2 [for $\ln(A_t/A_{t-1})$] = .65	

coefficients are negative, as predicted, although their impacts appear to be quite small. A 1% change in the growth rate of the military leads to a combined and immediate change of .05% in the prison population growth rate. Even the asymptotic effect is only .07%. However, the findings are more important than might first appear. The size of the military varies enormously in response to periods of war and peace. While the growth-rate mean is only .03, the standard deviation is .50. Moreover, the growth rate varies from a low of approximately −2.60 to a high of approximately 2.04. In practical terms, therefore, wartime mobilization leads to about a 14% decline in the prison-population growth rate, and demobilization leads to a 5% increase. Finally, since the military data are for the United States as a whole, the two elasticities are perhaps attenuated. For example, California did not participate in the Civil War and, for that period at least, the national figures are quite misleading.

Periods of economic depression also have a statistically significant effect. Both t-values are in excess of 1.64 (for lags of one year and two years), and the signs are positive, as predicted. The combined immediate effect is .07, while the combined asymptotic effect is approximately .10. However, since depressions are captured by a binary variable, an elasticity interpretation is not appropriate. Rather, the intercept (implicitly, zero) is shifted by an amount equal to "e" raised to the .07 power. This translates into an immediate effect of 1.07 and an asymptotic effect of 1.10. Within two years of the onset of depressions, the prison-population growth rate is multiplied by a factor of 1.07 (107%). Within four years, the (asymptotic) factor is 1.10 (110%).

Yet there is more to the story. Since the binary variable for depressions has been differenced, the 1–0 form has been translated into positive and negative spikes (of +1 and −1) around a base of zeros. Thus, the causal effect reflects spikes in the prison-population growth rate. However, spikes in the growth rate produce more lengthy shifts in level for the number of people imprisoned; a higher "plateau" materializes during depressions if the number of people imprisoned is considered.

The advent of the California Youth Authority was the only policy dummy variable retained, and it has a statistically significant (t = 2.97) positive impact on the State's prison-population growth rate.

With a regression coefficient of .16, the asymptotic effect is .24. Thus, the immediate result is that the prison population growth rate is multiplied by a factor of 1.17, and within three years (or so) it is multiplied by a factor of 1.27. And, as with depressions, spikes in the growth rate translate into shifts in level when the number of people in prison is considered. Unfortunately, it is not at all clear why the dummy variable for the California Youth Authority should be statistically significant when so many policy variables that were equally plausible on a priori grounds prove unimportant. The introduction of probation and indeterminant sentencing, for example, were certainly nontrivial innovations, yet no effects surfaced. We are left with the possibility of type I error, or a patently post hoc explanation, perhaps based on the well-known finding that diversifying the means of punishment often leads to a broader definition of who is subjected to sanctions.

The implementation of New Deal programs does not seem to affect the prison admissions growth rate. We had expected to find a decline in the growth rate; we assumed public works programs would increase the demand for labor. With an increase in the demand for labor, a reduction in the growth rate of the prison population should follow. Possibly, the public works projects were insufficient to substantially alter the demand for labor.

Perhaps more surprising is that the growth rate in the number of prisons also has no impact. Since prisons are added in discrete increments, we expected to find relatively discrete increments in the prison-population growth rate. However, the number of prisons is a rather indirect proxy for prison capacity, given the "creative" ways in which capacity can be defined; and new prisons may be used in the short run more to relieve overcrowding than to put larger numbers of new offenders behind bars. For example, in the year that the Folsom prison was completed (1881), 253 inmates were transferred from San Quentin to Folsom. This figure represents about 20% of the total number of individuals incarcerated in State facilities at that time, and illustrates exactly the sort of practice that makes the impact of prison capacity difficult to isolate. Indeed, new prisons may be built in response to social and political definitions of overcrowding that, in turn, are fueled by the kinds of variables we have already discussed. Then the feedback effects of prison capacity on punishment growth rates are spread across many years, are highly confounded with other causal factors, and are diluted by alternatives to incarceration (e.g., parole). In short, there is an interesting story to tell that future work should address.

Finally, one might be interested in how well the equation's predicted values track the observed values. Figure 3 shows a graph of this relationship, and the observed number of people in prison corresponds quite well to the predicted number of people in prison. The excellent fit is no doubt substantially due to the lagged value for the number of people in prison implicit in the growth-rate formulation (with a coefficient constrained to 1.0) and the lagged dependent variable specified in the model (with an unconstrained coefficient).

Consistent with our earlier analysis of the equilibrating process, the addition of the parole population makes little difference. The only new finding is that, not surprisingly, the prison- and parole-population growth rate increases with the "invention" of parole. Again, the fit is excellent by virtually any criterion.

Table 5 shows the results when punishment is defined in terms of the number of people admitted to state prisons for offenses not committed while on parole. While only 29% of the variance in the growth rate form is explained, the predicted and observed time paths for the number of individuals admitted correspond reasonably well. The important message is that for most of the series, there is no evidence of systematic errors in the model's fit.

The model for the growth rate of admissions is a bit different from the two earlier models we have been discussing. In particular, there is no need to include a lagged value of the admissions growth rate. This means that the growth rate of admissions (in contrast to the prison-population growth rate) responds quite rapidly to exogenous factors, and there is no need to

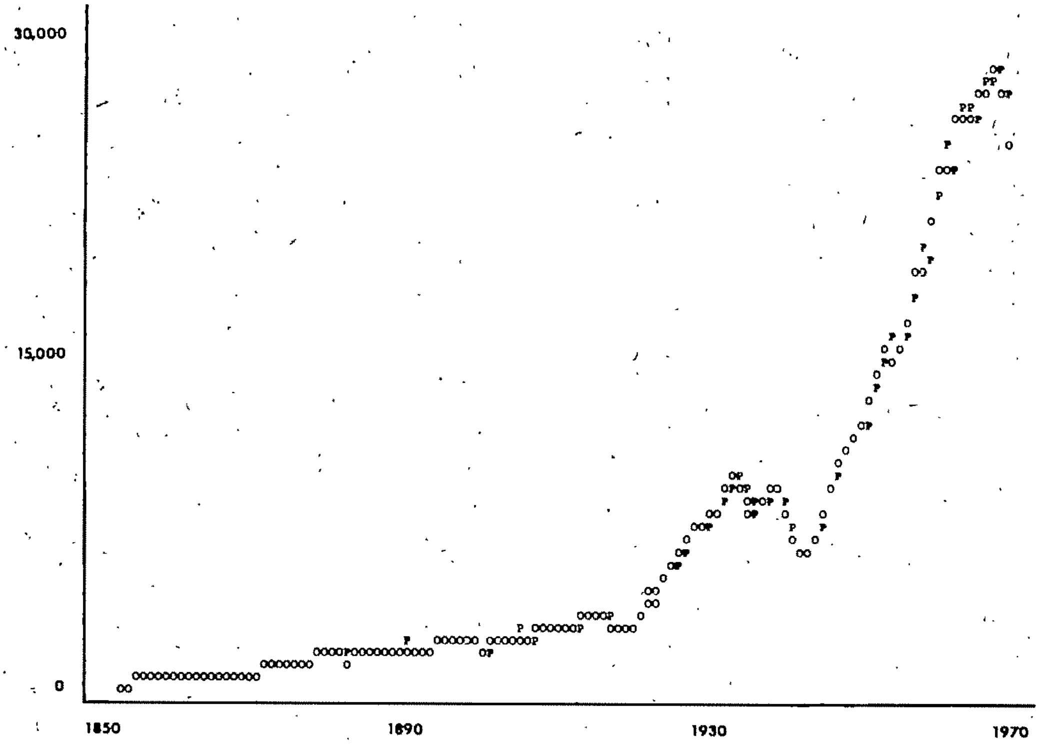


Figure 3. Observed (O) and Predicted (P) Values for Prison Population, 1853–1970
NOTE: In time periods where an O appears without a P, the observed and predicted values overlap.

consider asymptotic effects. Also, only one depression dummy variable is required, which is consistent with a more responsive endogenous variable.

The remaining effects shown in Table 5 are much as before. The admissions growth rate is increased with increases in the growth rate for the state's population

Table 5. Estimated Explanatory Model for "New" Prison Admissions

$$\ln (A_t/A_{t-1}) = \alpha_1 \ln (Pop_t/Pop_{t-1}) + \alpha_2 \ln (M_t/M_{t-1}) + \alpha_3 \ln (M_{t-1}/M_{t-2}) + \alpha_4 \ln (Pris_t/Pris_{t-1}) + \alpha_5(1 - B)CYA_t + \alpha_6(1 - B)D_{t-1} + \alpha_7(1 - B)ND_t + \ln a_t$$

Where:

A_t = Total prison admissions
Pop_t = California population
M_t = Number of people in the military
$Pris_t$ = Number of prisons
CYA_t = California Youth Authority dummy
D_t = Depression dummy
ND_t = New Deal dummy
a_t = White noise error

Parameter	Coefficient	t-Value
α_1 (Growth rate of California population)	.74	2.71
α_2 (Growth rate of the military, t)	−.06	−2.14
α_3 (Growth rate of the military, t−1)	−.07	−2.77
α_4 (Growth rate of prisons)	−.14	−0.92
α_5 (California Youth Authority)	.26	1.83
α_6 (Depression, t−1)	.17	3.94
α_7 (New Deal)	−.13	−1.26

R^2 [for $\ln A_t$] = .99 R^2 [for $\ln(A_t/A_{t-1})$] = .29

and with periods of depression, and decreased with increases in the growth rate for the number of people in the military (for the United States as a whole). Both the signs and the magnitude of the effects essentially replicate our earlier results.

Little changes when punishment is defined in terms of our broader definition of admissions. Despite a low R^2 of .29 in the growth rate form, the observed and predicted values for the number of admissions track very well. Perhaps the only issue worth discussing is the failure of the parole binary variable. Since our measurement of punishment includes individuals who committed new offenses on parole, the introduction of parole was expected to make a difference. However, the t-value of 1.48 falls short of statistical significance. We are inclined to dismiss this result as type II error, especially since, under some quite similar specifications, the t-value exceeds 1.64.

In summary, the results are much the same regardless of which measure of punishment is used. Growth rates of punishment respond primarily to major demographic trends, wars, and depressions. In this large sweep of history, changes in corrections policy appear to have almost no impact. Of course, a different story might have emerged had punishment been defined in other ways. Future research might do well to consider a wide variety of alternative definitions.

CONCLUSIONS

Our analyses of historical patterns of punishment in California have theoretical, statistical, and substantive implications. From a theoretical perspective, perhaps the major conclusion is that it is possible and desirable to formally model equilibrating processes when they are relevant to one's approach toward empirical phenomena. Past research on corrections has too often neglected the ways in which equilibration may occur; equilibria have either been assumed within a comparative static framework or ignored altogether. At the very least, this makes the existence of an equilibrium a nonfalsifiable premise and, in addition, overlooks the dynamic relationships required to properly understand longitudinal events. We suspect, therefore, that the range of formal models introduced above has important implications not only for studies of corrections, but for a wide variety of other sociological problems in which equilibrating processes are involved.

Besides directly representing a particular kind of equilibrating process, our dynamic models capture causal relationships that include short-term deviations from the steady-state time path, discrete shifts in the steady-state time path, and "continuous" shifts in the steady-state time path. The potential role of such factors in sociological formulations cannot be overemphasized. Simply put, social change is articulated through a number of complementary parameterizations that can be mixed and matched as needed. In particular, one is not limited to either "smooth" or "lumpy" approaches to social change.

Finally, the existence of an equilibrium within a dynamic formulation requires one or more definitions of a steady state. We have provided one common definition based on a constant rate of change. Whatever one's preferences are, it is clear that sociological models in which equilibration plays an important role must consider what it means for the system to be "at rest."

There are also important statistical lessons to be learned. Perhaps most important, along with Hendry and Mizon (1978) and Davidson et al. (1978), we have employed a specification strategy in which serial correlation in the residuals is not treated as a mere nuisance. Serially correlated residuals are often the result of specification errors that should not be ignored. In addition, even when ARIMA noise models can be used in certain instances to make proper adjustments (Hendry and Mizon, 1978), the resulting paramaterization is not likely to be substantively instructive.[12] One convenient

[12] To take a simple example, consider the common practice of first differencing (at a lag of 1) the endogenous variable in ARIMA models. This operation is the same as "regressing" Y_t on Y_{t-1} with the regression coefficient constrained to equal 1.0, and will eliminate any shifts in level (changes in intercept) in the Y_t time series. However, if the shifts in level are

outcome is that ordinary least-squares procedures will typically suffice, even with lagged values of the endogenous variable (i.e., consistent estimates are produced).

One major substantive conclusion is that when an equilibrating process is specified consistent with the traditional stability of punishment hypothesis, "Blumstein's constant" fails to materialize; either an F-test rejects the entire model or a t-test rejects the social control term. Equally important, there is solid evidence that the model on which these tests rest is fundamentally misspecified such that discrete and continuous shifts in the social control target must be considered. A second major conclusion, however, is that we also reject more general equilibrating models within the spirit of the Durkheim-Blumstein perspective. Even if one allows for discrete and/or continuous shifts in the steady-state target, equilibration does not surface. On the other hand, one might object to this conclusion on the grounds that, during the 120 years covered by our study, California experienced a number of significant social upheavals. For Blumstein (like Durkheim), a stable level of punishment was conditional upon a general stability in society.

We offer three rejoinders. First, none of the traditional stability of punishment statements we have been able to unearth define what qualifies as a major social upheaval. While events like the French Revolution are no doubt included, the role of wars, depressions, and demographic shifts within ongoing political institutions is unclear. Moreover, it is not apparent whether social upheavals need be discrete, dramatic events, or whether gradual transitions such as industrialization are also relevant.

Second, if one chooses to define dramatic social upheavals in terms of wars, depressions, industrialization, and the like, it may be difficult to find any stable intervals over the past 200 years, at least for Western societies. For our California data, the 1920s may be the only decade that qualifies.

Third and most important, our more flexible formulations (i.e., consistent with Models II and III) explicitly include the potential impact of major social upheavals; one can judge whether the social control ratio is part of an equilibrating process despite social disruptions. We are able to control for dramatic, exogenous shocks to the system and determine if a steady-state target(s) exists nevertheless. Thus, the Durkheim-Blumstein perspective is given the benefit of the doubt and still found wanting.

While we can find no evidence of an equilibrating process within the Durkheim-Blumstein perspective, punishment clearly responds to a number of important causal factors. Demographic trends and major historical events alter punishment growth rates in a substantial fashion. In contrast, penal reforms typically have little effect. These findings take on additional relevance in the current mood surrounding corrections policy. Some predict that the return of many states to determinate and/or prescribed sentencing will dramatically increase prison admissions and prison populations. However, in the longer sweep of history (perhaps less than a decade), recent reforms may well be swamped by larger social forces. As the post–World War II birth cohort matures, for example, there is good reason to expect that crime rates and (somewhat later) incarceration rates will substantially decline (Blumstein, Cohen, and Miller, 1980) almost regardless of sentencing practices.

At the same time, an enormous number of important questions remain unanswered. First, the role of the population growth rate needs further clarification. Since the impact was felt despite very short lags, it is difficult to argue that the growth rate in the population reflects something specific about any likely population at risk (with the possible exception of recent immigrants). Offenders are rarely eligible for prison sentences until their late teens, which means that, if the population growth rate was really a measure of potential-offender cohorts, much

really the result of one or more exogenous variables, no representation of their causal impact will be forthcoming. That is, the model "fits," but is not substantively meaningful.

longer lags should have surfaced (e.g., at 16 years). We could find no evidence for population effects at such lags. We are left with the somewhat unsatisfying conclusion that the population growth rate says something about the scale and complexity of society and the associated rate of punishment that apparently does not correspond to a simple input-output model.

Second, other definitions of punishment need to be considered. While the availability of appropriate data for long historical periods will certainly prove to be a significant constraint, many fruitful possibilities remain. For example, figures for the number of people on probation and in jails over time should be sought.

Third, future research should distinguish between the corrections experiences of men and women. If there is one conclusion from the recent work on women and crime (e.g., Adler, 1975; Simon, 1975; McCarthy, 1979; Bartel, 1979; Datesman and Scarpitti, 1980; Jurik, 1980; Freedman, 1981), it is that men and women do not make up a single, homogeneous offender population. Indeed, we have begun some work with our California data that confirms that models for female admissions are somewhat different from models for male admissions.

Fourth, a great deal will be learned if some efforts are directed toward unraveling the intervening mechanisms that we have ignored. In particular, significant insights would no doubt be gained if measures of crime and conviction rates could be obtained. With such measures in hand, one would be able to determine, for instance, if the impact of depressions on the growth rate of punishment operates through higher crime and conviction rates or is relatively independent of crime and conviction rates. Perhaps the major limitation of such efforts is that a lack of data before the turn of the century will place significant constraints on the length of the time-series data set.

Fifth, while we have made a strong case that the policy variables we have considered have almost no impact on the growth rate of punishment (despite an examination of a wide variety of lags), our list of policy alternations is hardly exhaustive. It may well turn out that policy effects will be found for policy changes we have inadvertently neglected.

Finally and more generally, we return to the question of equilibration. While we can find no evidence for equilibration within the Durkheim-Blumstein perspective, other equilibrating processes may well be at work. For example, one might redefine punishment in terms of its monetary costs and develop an equilibrating model with the growth rate in punishment costs as the endogenous variable and the proportion of tax dollars allocated to corrections as the steady-state (though perhaps shifting) target. Moreover, one may find it necessary to abandon a single-equation framework and build models closer to multiple-equation formulations describing regional economies. In short, a failure to find support for the Durkheim-Blumstein approach does not rule out other kinds of equilibrating processes related to punishment.

REFERENCES

Adler, Freda
1975 Sisters in Crime. New York: McGraw-Hill.

Bartel, Ann P.
1979 "Women and Crime: An Economic Analysis." Economic Inquiry 17:29–51.

Becker, Gary S.
1968 "Crime and Punishment: An Economic Approach." Journal of Political Economy 76:169–217.

Belsley, David A., Edwin Kuh, and Roy E. Welsch
1980 Regression Diagnostics. New York: John Wiley.

Berk, Richard A., Peter H. Rossi, and Kenneth J. Lennihan
1980a "Crime and poverty: some experimental evidence from ex-offenders." American Sociological Review 45:766–86.

Berk, Richard A., Donileen R. Loseke, Sarah F. Berk, and David Rauma
1980b "Bringing the cops back in: a study of efforts to make the criminal justice system more responsive to incidents of family violence." Social Science Research 9:193–215.

Blumstein, Alfred and Jacqueline Cohen
1973 "A theory of the stability of punishment." Journal of Criminal Law and Criminology 64:198–207.

Blumstein, Alfred, Jacqueline Cohen, and H. D. Miller
1980 "Demographically disaggregated projections of prison populations." Journal of Criminal Justice 8:1–26.

Blumstein, Alfred, Jacqueline Cohen, and Daniel Nagin
1977 "The dynamics of a homeostatic punish-

ment process." Journal of Criminal Law and Criminology 67:317–34.

Blumstein, Alfred and Soumyo Moitra
1979 "An analysis of the time series of the imprisonment rates in the states of the United States." Journal of Criminal Law and Criminology 70:376–90.

Bowker, Lee H.
1978 Women, Crime, and the Criminal Justice System. Lexington: D.C. Heath.

Box, George E. P., and Gwilym M. Jenkins
1976 Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.

Brenner, Harvey
1976 Estimating the Social Costs of National Economic Policy (for the Joint Economic Committee). Washington, D.C.: U.S. Government Printing Office.

Cahalan, Margaret
1979 "Trends in incarceration in the United States since 1880." Crime and Delinquency 25:9–41.

California State Board of Prison Directors
1920 Biennial Report. San Quentin: San Quentin Press.
1924 Biennial Report. San Quentin: San Quentin Press.

Chiang, Alpha C.
1974 Fundamental Methods of Mathematical Economics. New York: McGraw Hill.

Cohen, Lawrence E. and Marcus Felson
1979 "Social change and crime rate trends: a routine activity approach." American Sociological Review 44:588–607.

Cohen, Lawrence E., Marcus Felson, and Kenneth C. Land
1980 "Property crime rates in the United States: a macrodynamic analysis, 1947–1977; with ex ante forecasts for the mid-1980s." American Journal of Sociology 86:90–118.

Datesman, Susan K. and Frank R. Scarpitti
1980 Women, Crime, and Justice. New York: Oxford.

Davidson, James E. H., David F. Hendry, Frank Srba, and Stephen Yeo
1978 "Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom." The Economic Journal 88:661–92.

Durkheim, Emile
1964a The Division of Labor in Society. New York: Free Press.
1964b The Rules of Sociological Method. New York: Free Press.

Foucault, Michel
1979 Discipline and Punish: The Birth of the Prison. New York: Vintage Books.

Freedman, Estelle B.
1981 Their Sisters' Keepers: Women's Prison Reform in America, 1830–1930. Ann Arbor: University of Michigan Press.

Greenberg, David F.
1978 Mathematical Criminology. New Brunswick: Rutgers University Press.

Hendry, David F. and Graham E. Mizon
1978 "Serial correlation as a convenient simplification, not a nuisance: a comment on a study of the demand for money by the Bank of England." The Economic Journal 88:549–63.

Ignatieff, Michael
1978 A Just Measure of Pain: The Penitentiary in the Industrial Revolution, 1750–1850. New York: Columbia University Press.

Jenkins, Gwilym M.
1979 Practical Experiences with Modelling and Forecasting Time Series. St. Helier: Gwilym Jenkins & Partners.

Jurik, Nancy C.
1980 "Women ex-offenders in the TARP experiment." Pp. 319–34 in Peter H. Rossi, Richard A. Berk, and Kenneth J. Lenihan (eds.), Money, Work, and Crime. New York: Academic Press.

Lamott, Kenneth
1961 Chronicles of San Quentin. New York: David McKay.

Land, Kenneth C. and Marcus Felson
1976 "A general framework for building dynamic macro social indicator models: including an analysis of changes in crime rates and police expenditures." American Journal of Sociology 82:565–604.

McCarthy, Belinda Rodgers
1979 Easy Time. Lexington: D.C. Heath.

Melossi, Dario
1980 "Strategies of social control in capitalism: a comment on recent work." Contemporary Crisis 4:381–402.

Miller, Dan
1980 "Alternatives to incarceration: from total institutions to total systems." Ph.D. dissertation. Department of Sociology, University of California at Berkeley.

Pindyck, Robert S. and Daniel L. Rubinfeld
1981 Econometric Models and Economic Forecasts (2nd Edition). New York: McGraw Hill.

Perry, Mary Elizabeth
1980 Crime and Society in Early Modern Seville. Hanover: University Press of New England.

Rauma, David
forthcoming "Crime and punishment reconsidered: some comments on Blumstein's stability of punishment hypothesis." Journal of Criminal Law and Criminology.

Ruggiero, Guido
1980 Violence in Early Renaissance Venice. New Brunswick: Rutgers.

Rusche, Georg and Otto Kirchheimer
1968 Punishment and Social Structure. New York: Russell & Russell.

Simon, Rita James
1975 Women and Crime. Lexington: D.C. Heath.

Smart, Carol
1976 Women, Crime and Criminology: A Feminist Critique. London: Routledge & Kegan Paul.

Stigler, George J.
1970 "The optimum enforcement of laws." Journal of Political Economy 78:526–36.

Theil, Henri
1978 Introduction to Econometrics. Englewood Cliffs: Prentice-Hall.
Votey, Harold L., Jr. and Llad Phillips
1973 "Social goals and appropriate policy for corrections: an economic appraisal." Journal of Criminal Justice 1:219–40.
Waller, Irvin and Janet Chan
1974 "Prison use: a Canadian and international comparison." Canadian Law Quarterly 17:47–71.
Zellner, Arnold and Franz Palm
1974 "Time series analysis and simultaneous equation econometric models." Journal of Econometrics 2:17–54.
Yaley, Barbara Jeanne
1980 "Habits of industry: labor and penal policy in California, 1849–1940." Ph.D. dissertation. Department of Sociology, University of California at Santa Cruz.

THE JAPANESE AMERICANS: CHANGING PATTERNS OF ASSIMILATION OVER THREE GENERATIONS*

DARREL MONTERO
Arizona State University

Using a national sample of three generations of Japanese Americans, we examine the question of whether socioeconomic mobility leads to cultural, structural, and marital assimilation. On three of our four measures of assimilation, the higher the socioeconomic achievement, the greater the assimilation. These findings are discussed in light of the substantial socioeconomic gains the Japanese American community is presently making and the effect of those gains upon the assimilation of future generations of Japanese Americans. Finally, we develop a sociohistorical model that attempts to explain the unfolding of the assimilation process for our three-generational sample of Japanese Americans.

An important issue regarding the study of assimilation involves the question of generational differences. Robert Park (1914, 1950) and others have suggested that assimilation unfolds naturally across generations (Dohrenwend and Smith, 1962; Gordon, 1964, 1975; Taft, 1957). Other writers have predicted a pattern of ethnic revival (Greeley, 1971; Hansen, 1937, 1952; Lyman and Douglass, 1973). Hansen (1937, 1952), for example, terms his model the "Law of the Return of the Third Generation." He suggests that the third generation, more secure in its socioeconomic status and American identity, becomes interested in the ethnic heritage that the second generation neglected in its efforts to overcome discrimination and marginality. Simply stated, "What the child wishes to forget, the grandchild wishes to remember." While Hansen's thesis (1937, 1952) has received some support (Herberg, 1955; Kennedy, 1944, 1952), it has numerous critics as well (Abramson, 1973; Goering, 1971; Levy and Kramer, 1973; Rose, 1976).

Most studies of assimilation report results based upon cross-sectional samples of ethnic communities, thereby obscuring the intrafamilial generational changes in assimilation (e.g., Caudill, 1952; Feagin

* Direct all correspondence to Darrel Montero, West Hall, Arizona State University, Tempe, Arizona 85281.

The study was made possible through grants from the Japanese American Citizens League, the Carnegie Corporation of New York, and the National Institute of Mental Health (Grant No. 5 RO1 MH12789-04). Computing assistance was obtained from the Health Sciences Computing Facility, UCLA, sponsored by NIH Special Research Resources Grant RR-3. I am especially appreciative of the generous assistance and extensive comments provided by Gene N. Levine, the principal investigator of the Japanese American Research Project, Judith McDowell, and Tara McLaughlin. Also I gratefully acknowledge the valuable comments made on earlier drafts of this paper by Kenneth D. Bailey, Edna Bonacich, Marvin Dawkins, Stanford M. Lyman, Tara McLaughlin, Patricia Rathbun, R. A. Schermerhorn, Frank Stratford, Paul Wong, and two anonymous referees for the journal. Computer assistance was provided by Chris Turner. Technical assistance was provided by Betty Wood.

and Fujitaki, 1972). The central concern of this study is to examine generational differences in assimilation among Japanese Americans, especially whether socioeconomic mobility leads to cultural, structural, and marital assimilation among the second generation. Following Gordon (1964:71), cultural assimilation refers to the process characterized by a group changing its behavioral patterns (e.g., religion) to those of the host society. Structural assimilation includes large-scale entrance into cliques, clubs, and institutions of the host society on a primary-group level. Marital assimilation refers to large-scale intermarriage between members of the host society and the immigrant group. Our findings will help determine whether Park's or Hansen's thesis is more viable.

Ianni (1957) observed that most studies of the assimilation of American ethnic groups have revealed some positive relationship between social mobility and assimilation. While few of these studies measured social mobility, they usually contained observations by field workers that ethnic group members with higher social status tended to be more assimilated (Spiro, 1955). In Yankee City, for example, the least assimilated ethnic groups and the least assimilated members within the various groups were also characterized by low social status (Warner and Srole, 1945). Similar relationships have been described for Greeks (Fairchild, 1911), the Irish (Wittke, 1956), the Italians (Child, 1943; Ianni, 1952; Myers, 1949; Whyte, 1943), and the Jews (Gordon, 1949).

Japanese immigrants and their descendants in the Unted States have achieved high socioeconomic mobility. Japanese Americans have the highest median number of years of education completed (12.5 years) of all racial groups (U.S. Bureau of the Census, 1973a, 1973b, 1973c, 1973d) and are nearly twice as likely to be employed as professionals. Does socioeconomic status affect assimilation? Does the rate of assimilation increase for each subsequent generation, or do we witness an ethnic renaissance among the third generation, as predicted by some writers (Hansen, 1937, 1952)? We approach these questions with data from a three-generational, national sample of Japanese Americans focusing particularly upon the Nisei.

METHOD

Our major independent variables are education and occupation; the dependent variables include four indicators of assimilation: visiting patterns with relatives, ethnicity of favorite organization, ethnicity of closest friends, and ethnicity of spouse.

The Sample. The data are derived from a sample of 2,304 second-generation and 802 third-generation Japanese Americans. This sample is part of a larger sociohistorical, three-generational study of continental Japanese Americans conducted by the Japanese American Research Project at UCLA. The purpose of the study was to sample and to interview surviving members of the first generation (the Issei) who had arrived in the continental United States before the 1924 exclusion act. In 1963 an attempt was made to list every Issei immigrant still living in the United States, excluding Hawaii and Alaska. The listing was largely derived from Japanese association membership lists and supplemented by lists from Buddhist and Christian churches; it totaled about 18,000 persons. Undoubtedly, these lists were incomplete, leaving out people who had failed to join an organization. The Japanese immigrant community, however, was one of the most highly organized of immigrant groups (Park and Miller, 1921:168), and the error is not nearly as great as it would be for some other immigrant nationalities.

During 1964–1966, interviews with a random sample of listed Issei were carried out by specially recruited and trained, bilingual Japanese American interviewers. With a respondent refusal rate of less than one percent, this yielded a sample of 1,047 Issei. Of these, 141 either had not married, had failed to produce offspring, or had had children none of whom consented to be interviewed or returned our questionnaires. Our effective multigenerational sample is thus 906 Issei. These Issei had 3,817 Nisei (or second-generation) off-

spring. The offspring were surveyed during 1967, in three stages, with a greatly modified version of the Issei interview schedule. Overall, we achieved a 60.5% response rate for the Nisei, yielding a total Nisei sample of 2,304 respondents. Of these, 34.4% (N = 792) were interviewed by non-Japanese, professional interviewers from the National Opinion Research Center (NORC) in areas that are regular NORC primary sampling units. Telephone interviews were conducted by NORC personnel when a desired respondent lived more than 50 miles beyond any NORC primary sampling unit (3.5% of the sample). The remainder (62.1%) were sent mail questionnaires concurrently. This latter method yielded a 49% response rate after three mailings. In late 1967, questionnaires were mailed to the Nisei's 1,063 adult offspring (18 years of age or older), the Sansei (third generation). Three mailings elicited returns from 75.5%, or 802 of the respondents.

Sample Comparisons. We examined the Japanese American population figures from the 1960 and 1970 censuses, and it is clear that the Issei far exceeded our listing of 18,000. However, of the 101,656 Issei enumerated in 1960 (which included those in Hawaii and Alaska), 43% were younger than thirty-five (U.S. Bureau of the Census, 1963:8) and could not possibly have immigrated prior to 1924. Furthermore, among those who were born in time to have been Issei by our definition, many must have traveled as adults and thus arrived well after 1924.

The sample consisted solely of families in which a member of the immigrant generation had survived into the mid-1960s. We had a somewhat higher concentration of later immigrants than the immigration records report. This bias had an effect on the Nisei sample that omitted the children of dead Issei: those in our second-generation sample were more likely to be the progeny of Issei who came in the later phases of the migration to the United States. Our sample was considerably more nearly representative of the Japanese immigrants to the mainland who persisted there to 1930, at what might be said to be the height of the Issei generation's vigor and development.

To gauge the extent of sample bias, we compared some of the characteristics of our Nisei sample with corresponding census data. An age-distribution comparison was not possible because the census does not divide the Japanese American population by generation, but we were able to compare regional dispersion. We found that while the sample Nisei were slightly more concentrated in the West than was the case for the total Nisei population, the difference was small.[1]

We were unable to interview Nisei and Sansei whose parents and grandparents died before the early 1960s or who had returned to Japan. To determine the extent of this potential sampling error, we examined death records at the two Japanese American mortuaries in Los Angeles. This search resulted in a subsample of 38 Nisei and 29 Sansei descendants and revealed no remarkable differences, except that, as expected, the survivors tended to be older than those Nisei respondents who had at least one parent alive.

We probably have too many Issei respondents who are immersed in organized Japanese American life and too few who are unaffiliated. The bias is inherent in the original Issei listing procedure. We have probably undersampled unaffiliated persons and taken a disproportionate share of the better educated and better-off. This sampling bias is probably smaller than for other immigrant groups because Japanese exhibit such a high degree of organizational affiliation (Park and Miller, 1921:168).

FINDINGS

About 38% of our respondents live in the metropolis of Japanese America, Los Angeles. Only 24% of the Nisei live away from the West Coast. Historically, there has been a fairly dramatic change in the dominant neighborhood ethnicity for the Japanese Americans. One-third of the Nisei living in 1915 reported that they resided in predominantly Japanese Ameri-

[1] Further comparisons between our sample and various census and immigration data are available from the author on request.

can neighborhoods, one-third in mixed neighborhoods, and one-third in mainly non-Japanese American neighborhoods.[2]

Since 1915, the proportion of Nisei living in predominantly Japanese American neighborhoods has been steadily declining. The most recent year for which our respondents reported (1967) indicates that the percentage of Nisei respondents residing in predominantly Japanese American neighborhoods has declined dramatically to 4%; the percentage residing in ethnically mixed neighborhoods has remained relatively stable at 38%, while those Nisei living in predominantly non-Japanese American neighborhoods has increased substantially to a majority of 58%.

This shift out of Japanese American ghettos would seem at first glance to mirror the demise of Japanese American ethnic communities as we have come to know them (Lyman, 1970). In his study of New Haven's Italian community, for example, Myers (1950) found that residential dispersion tends to be related to assimilation to the dominant social system. In a study of ten American cities, Lieberson (1961) found that residential segregation was related to ability to speak English, citizenship, intermarriage, and occupational composition.

How do ethnic ties of these "nonghetto" residents differ from the 42% who live in predominantly Japanese American and ethnically mixed neighborhoods? One-sixth of the Nisei do not live within the same city or county as one or more of their relatives. Nearly a majority live alone in the neighborhood without one or more of their relatives.

Since religion is an indicator of assimilation, we want to determine the effect of religious affiliation upon the presence of relatives in the city. We found that 25% of the Protestants, 23% of the Catholics, and only 9% of the Buddhists had no relatives living within the same city. Those Nisei most likely to be "alone" in their community were middle class Protestants (33%), Catholics (31%), and Buddhists (11%). The pattern was similar for working-class Nisei: 18% of Protestants, 16% of Catholics, and 6% of Buddhists were isolated in their community.

Lenski (1963:214–7) provides some comparative figures regarding those persons who had no relatives in Detroit, except their immediate family. His findings are consistent with our results: Protestants are the most likely to be alone in their community.

Sociological research makes it abundantly clear that spatial mobility facilitates vertical mobility (Barber, 1957: 418–21; Lieberson, 1961; Lipset and Bendix, 1959: 157–64; Merton, 1957: 400–1; Stuckert, 1963). If this is true, people whose ties with kin bind them to their community of birth are at a disadvantage in the competition for advancement. This factor may well contribute to the different rates of mobility within our sample of Japanese Americans.

Nisei Community Interaction and Affiliation

Among Nisei who have relatives in the same metropolitan area, 29% choose not to pay them regular visits, while 44% visit or are visited by relatives at least once a week. Thus presence of relatives in the same metropolitan area does not automatically lead to intensive visiting patterns.

Only 45% of our respondents belong to a Japanese American organization; of those who belong to any organization, 55% devote the most time to a non-Japanese organization. Although the Issei almost uniformly belonged to only Japanese organizations, less than half (45%) of the Nisei and less than one-third (31%) of the Sansei were members of a Japanese organization. Lopata's (1976:109) study of Polish American organizational affiliation revealed the same general pattern.

Finally, our data reveal that 53% of all Nisei and 74% percent of the Sansei claim that one or both of their two closest friends are non-Japanese. In a study of Mexican Americans in three cities,

[2] Lieberson (1973:562–3) reminds us that we must be sensitive to the impact of generational effects and further emphasizes that the absence of a generational control may create misleading impressions of the causes and nature of a group's changes over time. Therefore, we examine each of the three generations separately.

Table 1. Nisei Occupation and Education by Indicators of Assimilation

	Visit Relatives Infrequently[a]		Ethnicity of Favorite Organization is Non-Japanese		Ethnicity of One or Both of Two Closest Friends is Non-Japanese		Ethnicity of Spouse is Non-Japanese	
Occupation (males only)								
Professional	39%	(359)	70%	(278)	64%	(387)	16%	(289)
Proprietor	22	(195)	48	(146)	47	(203)	6	(182)
Clerical	24	(124)	44	(82)	47	(137)	9	(95)
Blue collar	21	(139)	43	(86)	45	(143)	10	(111)
Service	21	(100)	19	(83)	31	(109)	8	(89)
Farm	21	(132)	43	(122)	39	(153)	5	(131)
Education								
Postgraduate	45%	(280)	79%	(220)	73%	(299)	22%	(221)
College graduate	37	(281)	63	(214)	56	(312)	13	(234)
Some college	25	(635)	54	(407)	53	(678)	9	(511)
High-school graduate	23	(864)	37	(579)	45	(945)	6	(797)

[a] Visiting relatives is derived from the question: "About how many times in the past month have you visited or been visited by relatives living in the same metropolitan area as you?" No visits that month was defined as infrequent.

ethnicity of present friends ranged from a high of 55% having all Mexicans as their present friends to a low of 22% (Moore, 1976:137).

Evidence in terms of neighborhood ethnicity, visiting patterns, organizational membership, and friendship patterns suggest a movement toward assimilation. Who are these assimilated Nisei? Are they outcasts, or are they willing converts to another way of life? Is their action purposeful or accidental? Is socioeconomic status positively related to assimilation?

Socioeconomic status and visiting patterns. Nisei professionals are the least likely to visit their relatives.[3] This finding supports our hypothesis that socioeconomic status is positively related to assimilation.[4] Education is also positively related to assimilation (Table 1). Nisei with postgraduate training are the least likely to visit relatives (45%). Lenski (1963:214) has observed that one of the best indicators of the importance attached to family and kin group by modern Americans is their willingness, or unwillingness, to leave their native community and migrate elsewhere. Migration of this type generally involves a physical separation of the individual from those relatives with whom he has the closest ties. A person usually migrates in response to the lure of economic or vocational opportunities; we may regard migration as an indicator of the importance the person attaches to the kin group when its ties compete directly with the prospect of more money or a better job (Schneider and Homans, 1955).

Socioeconomic status and ethnicity of favorite organization. We asked the Nisei about the ethnic make-up of the organization to which they devote the bulk of their leisure time. We reasoned that should the Nisei spend more of their leisure time in a non–Japanese American organization, this would indicate greater structural assimilation. We found that occupation is positively related to structural assimilation. Professional occupational status often brings with it other pressures and commitments that intervene and may take priority over local, family, and ethnic concerns. We found that education is also positively related to structural assimilation. Seventy-nine percent of the postgraduate educated Nisei chose a non–Japanese American organization as their favorite.

Socioeconomic status and friends' ethnicity. Our next measure of assimilation is the ethnicity of the two persons described by the respondents as closest friends. These data should provide a good indicator of the Nisei rate of structural assimilation. We expected that occupation would be positively related to this mea-

[3] In examining the Nisei visiting patterns, we use as our base only those respondents who report having relatives living within the metropolitan area.

[4] We collected occupational data only for the male Nisei.

sure of assimilation, and the data in Table 1 support this hypothesis. Professional Nisei are more than twice as likely as service workers to have non-Japanese friends. Nisei professionals have greater contact with the larger society—first, through the experience of higher education, and second, through the predominantly Anglo populated professions, with the consequent push and pull of professional demands and associations.

We find that education is positively related to non-Japanese friendship choices: the higher the educational attainment, the more likely Nisei are to have one or more non-Japanese friends. Seventy-three percent of the postgraduate as compared with only 45% of the high school educated Nisei report interethnic friendships.

Socioeconomic status and spouse ethnicity. Social scientists generally agree that one of the most telling indicators of the degree of assimilation of one ethnic group into another is the extent of outmarriage from the immigrant to host society (Barron, 1972; Gordon, 1964; Merton, 1941; Tinker, 1973). Although the outmarriage rate for the immigrant Issei was less than 1%, it increased to 10% of the Nisei, and to 40% of the Sansei.[5] While the census data (U.S. Bureau of the Census 1970:262) do not provide overall rates of intermarriage, they do provide rates by sex for those who outmarried. These data indicate that white females outmarry at the rate of .3% and white males choose other than a white spouse .4% of the time. American Indians have the highest rate of intermarriage: 39% of the women outmarry, and 35.8% of the men marry out of their ethnic group. For Blacks the percentage is .7 for women and 1.5 for men; for Chinese it is 12.2 and 13.5; Filipinos, 27.2 and 33.5; Mexicans, 16.7 and 16.1; and Puerto Ricans, 18.2 and 19.4. The rate for Japanese women is 33.2%, and 11.4 for men.

Clearly, if we find that socioeconomic status is postively related to outmarriage, given the substantial educational and occupational gains the Nisei and the Sansei are making, this would lead us to predict an accelerating rate of assimilation among the Japanese Americans in the future (Table 1).

We encounter an irregular relationship between occupation and spouse ethnicity. Although the Nisei professionals have the highest outmarriage rate—one in six—and the farmers have the lowest—one in 20—little else is consistent. The findings at the extremes are clearly related to mate-selection opportunity, education, and style of living. The daily experiences of the Nisei professional would more often bring him or her into contact with non-Japanese, the reverse being true of the Nisei farmer.

We found that education is positively related to exogamy. As indicated earlier, this may be largely a result of attendance at predominantly white college campuses for five or more years, with the concomitant internalization of new values and attitudes during that period of resocialization.

A multivariate analysis was conducted using our four dependent variables by occupation, controlling for education and age (Table 2). To maximize cell frequencies, as well as for clarity of presentation, we dichotomized both our independent and dependent variables. Our intent is to elaborate and specify the relationships reported in Table 1. Our findings tend to support our hypothesis. For example, when we look at social visiting, we find that regardless of age, the higher educated white-collar Nisei visit less frequently than their blue-collar counterparts. Regarding ethnicity of favorite organization, it is the white-collar Nisei who are more likely to claim a non-Japanese organization as their favorite. For ethnicity of closest friends, we find that regardless of age or education, white-collar Nisei are more likely to select non-Japanese as closest friends. Finally, for ethnicity of spouse, we find that the pattern is the same as the two-variable relationship: education is positively related to racial intermarriage while the pattern for occupation is inconsistent.

[5] These results, which differ markedly from those reported by the U.S. Bureau of the Census (1970:262), may in part be explained by the fact that we report data by generation, whereas the census data include all Japanese outmarriages regardless of age or generation. The data, therefore, need to be interpreted cautiously given that they include both war brides and Sansei intermarriage figures.

Table 2. Visiting Patterns, Ethnicity of Organization, Friends, and Spouse by Occupation, Controlling for Education and Age

	≤43 Years Old				≥44 Years Old			
	≤High School		> High School		≤High School		> High School	
	Blue Collar	White Collar	Blue Collar	White Collar	Blue Collar	White Collar	Blue Collar	White Collar
Social visiting infrequent	56%	45%	45%	48%	55%	47%	47%	53%
	(66)	(79)	(25)	(194)	(80)	(100)	(24)	(141)
Ethnicity of favorite organization	41	53	69	66	30	49	49	61
is non-Japanese	(37)	(50)	(31)	(271)	(40)	(94)	(23)	(196)
Ethnicity of closest friends	41	51	49	58	38	47	51	58
is non-Japanese	(67)	(131)	(40)	(388)	(77)	(131)	(34)	(235)
Ethnicity of spouse	9	8	22	15	5	4	11	8
is non-Japanese	(12)	(18)	(14)	(75)	(9)	(9)	(6)	(27)

NOTE: The table above presents the relationships between each of our four dependent variables by occupation, holding education and age constant. We also treated education and age as independent variables while holding the other demographic variables constant. The results of these analyses are consistent with those reported above as well as with the original two-variable relationships. Although space limitations prevent our reporting these results within the body of the present paper, these tables are available from the author upon request.

Assimilation over Three Generations

In order to have an overview of the differential rates of assimilation for the Japanese Americans, we present six indicators of cultural, structural, and marital assimilation by Issei, Nisei, and Sansei generations (Table 3). The trend toward greater assimilation for each succeeding generation is clear. Issei, Nisei, and Sansei rate increasingly higher on each of the indicators of assimilation. These findings do not lend support to Hansen's (1937, 1952) third-generation, ethnic-revival thesis. Rather than witnessing an ethnic revival among the Sansei, each indicator of assimilation points to a pattern which is more consistent with Park's (1950) concept of assimilation.

The Duality of Attraction Model. Although several authors have reported increasing rates of assimilation across generations (e.g., Goldstein and Goldscheider, 1968; Sandberg, 1974), few have attempted to explain exactly how and why the process occurs. We present below a model which seeks to explain the unfolding of this phenomenon. Our findings suggest a duality of attraction pulling Japanese Americans in two different directions: there is a desire to preserve one's sense of ethnicity in the face of the pull of socioeconomic advancement.

When the first-generation immigrants,

Table 3. Assimilation by Generation: Issei, Nisei, and Sansei

Types of Assimilation	Issei	Nisei	Sansei
Cultural Assimilation			
Non-Buddhist religion	35%	63%	76%
	(858)	(2229)	(786)
Structural Assimilation			
Ethnicity of favorite organization is non-Japanese	†	55%	69%
		(1422)	(218)
Non-Japanese neighborhood	45%	58%	67%
	(808)	(2295)	(801)
Ethnicity of one or both of closest friends is non-Japanese	†	53%	74%
		(2239)	(795)
Visiting with relatives infrequent	†	29%	23%
		(2064)	(695)
Marital Assimilation			
Ethnicity of spouse is non-Japanese	1%	10%	40%
	(906)	(1765)	(238)

† Data not available.

the Issei, arrived in the United States, they sought (and found) security within the ethnic enclave. Regarding themselves as sojourners, they came to the United States in search of financial security, and planned to return to their homeland once they made their fortunes (Miyamoto, 1939). They settled in ghettos, comforted by the similarity of language, custom, and culture. They organized according to prefecture of origin, and formed ethnic institutions as systems of mutual aid and support, such as rotating credit associations (Light, 1972).

An ethnic economy was formed, composed of small shopkeepers and proprietorships that served the Japanese American community. The ethnic economy served to hold together the Japanese American community, to reinforce values, customs, behavior, and social control (Bonacich, 1975).

The value system of the Japanese encouraged economic success, which is one reason why they have come to the United States (Kitano, 1976). Through hard work and perseverance (Petersen, 1970, 1971), aided as well by their value system, which was compatible with that of the American middle class (Kitano, 1976), they soon achieved a measure of financial security. This allowed them to send their children, the Nisei, to universities and other institutions of higher education.

Finding themselves on heterogeneous college campuses with few other Japanese students, the Nisei necessarily formed friendships with non-Japanese. Their educational achievements opened the door to higher occupational status than had been possible for their parents. The Japanese value system holds family and tradition very dear, and the Nisei were confronted with a difficult decision: a duality of attraction. They were feeling the lure of economic advancement from the host American society, while feeling the conflicting pull of accepting employment offered by the ethnic economy and thereby remaining within this supportive network. Those who have chosen to become doctors or lawyers, for instance, could choose to serve their ethnic community. On the other hand, those who chose to become engineers, teachers, and business executives found that advancement in those fields required geographical mobility.

Since geographic mobility is generally required to achieve economic advancement, this entailed a willingness to leave kith and kin to seek one's fortune. The tipping point, as it were, comes when the social-psychological value and comfort provided by the ethnic community begins to diminish relative to the lure of the socioeconomic advancement which higher education allows.

There may be no going back once this tipping point is reached. Furthermore, even where there is a desire to return to the ethnic fold, the socioeconomic advantages of "tipping over" makes it impossible to do anything but that. Originally, ethnic cohesion was an advantage in that it helped the Japanese to gain a foothold in American society. In contrast, affiliation with the ethnic community for the Nisei now becomes a relative disadvantage.

Economic success encourages movement from little Tokyos and Japan-towns into the surrounding suburbs, thus splintering community ties. The process of assimilation may be accelerated by the absence of the Issei, who have remained in the ethnic enclave and are no longer present to ensure that their grandchildren, the Sansei, receive a traditional Japanese upbringing. If the Sansei are reared in largely non–Japanese American communities, the absence of other Japanese children to serve as agents of socialization will further speed the assimilation process.

Another factor which contributes to their assimilation is that, unlike their parents, the Nisei do not consider themselves sojourners (Miyamoto, 1939). They have known only the United States as home and may wish consciously to assimilate.

Finally, as the Nisei succeed in occupations which require them to live in largely non-Japanese communities, they form friendships and professional acquaintances among Caucasians. These newly established relationships have important implications for racial intermarriage and ultimately for their further assimilation into the larger society.

A competing thesis proposed by Barton

(1975) suggests that, rather than assimilating into the larger society, ethnic groups may be breaking up into groups which are more varied, but still remain unassimilated. In the Italian community, this subgroup disintegration Barton describes went from Tuscan to Italian to Catholic.

The implication is that Japanese ethnic identity may be breaking apart in favor of an Asian American ethnic identity, so that a new solidarity is replacing the old. Wong (1972, 1981) and Uyematsu (1971) point out that over the last decade, with the development of the Asian American political movement, the changes in historical attitudes, and the greater numbers of Asians in metropolitan areas, Japanese are beginning to intermarry other Asian groups in larger numbers. Wong (1972, 1981) and Uyematsu (1971) acknowledge, however, that historically an Asian American identity as a social concept has not been viable.

Our Nisei data do not support Barton's (1975) Asian identity thesis. Regarding intermarriage, for example, when we asked the Issei who they preferred their children to marry, in order of preference, the following pattern emerged: Nisei, Japanese-born, Caucasians, Chinese, Koreans, Filipinos, Mexicans, and Blacks. Moreover, our data indicate that of those Nisei who have intermarried (10%) some 92% have married not other Asian Americans, but rather Caucasians. In the next generation, the Sansei, 40% have outmarried. Following their parents' pattern, a majority (78%) married Caucasians, 19% other Asians, and 3% other ethnic groups. More recent intermarriage data collected by Kikumura and Kitano (1973:72–3) and Tinker (1973) reflect a similar pattern. In addition, our findings regarding other indicators of assimilation–ethnicity of neighborhood, organization, and closest friends–point to a similar movement toward the larger Caucasian community.

SUMMARY AND CONCLUSIONS

We examined four basic indicators of assimilation for the Japanese American community in the United States: visiting patterns with relatives, ethnicity of two closest friends, ethnicity of favorite organization, and ethnicity of spouse.

Our major independent variables were occupation and education. Because occupational status in many instances goes hand in hand with level of education, we found, with one exception, that these two varibles were positively related to our indicators of assimilation. The exception to this pattern was ethnicity of spouse; while education is positively related to interracial marriage, occupation is not.

The trend toward greater assimilation is clear. Issei, Nisei, and Sansei rate increasingly higher on every indicator of assimilation. It is apparent that both the Nisei and Sansei are making remarkable strides in socioeconomic advancement, and that advancement in turn is positively related to assimilation.

Can the ethnic community in fact remain intact when its members are being scattered to the four winds as they seek to advance themselves socioeconomically? Our findings suggest that on every indicator of assimilation it is the socioeconomically successful Nisei who are the most assimilated and cut off from the ethnic community. Since the majority of our Nisei respondents are making considerable economic strides, this suggests an accelerating rate of assimilation for the Japanese American population as a whole. Ironically, that very assimilation may suggest the demise of some of the Japanese American community's traditional values, which were so instrumental in catapulting its members to these heights. The demise of these values in turn may serve to bring about the leveling off of Nisei and Sansei socioeconomic achievement. As their values become more congruent with the larger American society, Japanese Americans will most likely begin to mirror the achievement patterns of American society in general.

Given the dramatically increasing trend of outmarriage among the Sansei, with its concomitant erosion of ethnic ties and affiliation, we are justified in wondering whether a Japanese American ethnic community can be maintained into the next generation—the Yonsei. If it cannot,

the survival of other distinct ethnic groups may be similarly uncertain as their members advance socioeconomically.

REFERENCES

Abramson, Harold J.
1973 Ethnic Diversity in Catholic America. New York: Wiley.

Barber, Bernard
1957 Social Stratification. New York: Harcourt, Brace.

Barron, M.
1972 The Blending Americans, Patterns of Intermarriage. Chicago: Quadrangle.

Barton, Josef J.
1975 Peasants and Strangers: Italians, Rumanians, and Slovaks in an American City, 1890–1950. Cambridge, Mass.: Harvard University Press.

Bonacich, Edna
1975 "Small business and Japanese American ethnic solidarity." Amerasia Journal 3:96–112.

Caudill, William
1952 "Japanese-American personality and acculturation." Genetic Psychology Monographs 45:3–101.

Child, Irvin L.
1943 Italian or American: The Second Generation in Conflict. New Haven: Yale University Press.

Dohrenwend, Bruce P. and Robert J. Smith
1962 "Toward a theory of acculturation." Southwestern Journal of Anthropology 18:30–9.

Fairchild, Henry Pratt
1911 Greek Immigration to the United States. New Haven: Yale University Press.

Feagin, Joe R. and Nancy Fujitaki
1972 "On the assimilation of Japanese Americans." Amerasia Journal 1:13–30.

Goering, John M.
1971 "The emergence of ethnic interests: a case of serendipity." Social Forces 49:379–84.

Goldstein, Sidney and Calvin Goldscheider
1968 Jewish Americans: Three Generations in a Jewish Community. Englewood Cliffs, N.J.: Prentice-Hall.

Gordon, Albert I.
1949 Jews in Transition. Minneapolis: University of Minnesota Press.

Gordon, Milton
1964 Assimilation in American Life. New York: Oxford University Press.
1975 "Toward a general theory of racial and ethnic group relations." Pp. 84–110 in N. Glazer and D. P. Moynihan (eds.), Ethnicity. Cambridge: Harvard University Press.

Greeley, Andrew
1971 Why Can't They be Like Us? New York: Dutton.

Hansen, Marcus L.
1937 The Problem of the Third-Generation Immigrant. Rock Island, Ill.: Augustana Historical Society.
1952 "The third generation in America." Commentary 14:492–503.

Herberg, Will
1955 Protestant, Catholic, Jew. New York: Doubleday.

Ianni, Francis A. J.
1957 "Residential and occupational mobility as indices of the acculturation of an ethnic group." Social Forces 36:65–72.
1952 "The acculturation of the Italo-Americans in Norristown, Pennsylvania." Doctoral dissertation, Department of Sociology, Pennsylvania State University.

Kennedy, Ruby Jo Reeves
1944 "Single or triple melting pot? Intermarriage in New Haven, 1870–1940." American Journal of Sociology 49:331–9.
1952 "Single or triple melting pot? Intermarriage trends in New Haven, 1870–1950." American Journal of Sociology 58:56–9.

Kikumura, Akemi and Harry H. L. Kitano
1973 "Interracial marriage: a picture of the Japanese-Americans." Journal of Social Issues 29:67–81.

Kitano, H. L.
1976 Japanese Americans: The Evolution of a Subculture. Second Edition. Englewood Cliffs, New Jersey: Prentice-Hall.

Lenski, Gerhard
1963 The Religious Factor. Garden City, New York: Anchor Books.

Levy, Mark R. and Michael S. Kramer
1973 The Ethnic Factor. New York: Simon and Schuster.

Lieberson, Stanley
1961 "The impact of residential segregation on ethnic assimilation." Social Forces 40:52–7.
1973 "Generational differences among Blacks in the North." American Journal of Sociology 79:550–65.

Light, Ivan H.
1972 Ethnic Enterprise in America: Business and Welfare Among Chinese, Japanese, and Blacks. Berkeley and Los Angeles: University of California Press.

Lipset, Seymour M. and Reinhard Bendix
1959 Social Mobility in Industrial Society. Berkeley: University of California Press.

Lopata, Helena Znaniecki
1976 Polish Americans. Englewood Cliffs, New Jersey: Prentice-Hall.

Lyman, Stanford M.
1970 "Contrasts in the community organization of Chinese and Japanese in North America." Pp. 57–63 in S. M. Lyman (ed.), The Asian in the West. Social Science and Humanities Publication No. 4. Reno: Western Studies Center, Desert Research Institute, University of Nevada System.

Lyman, Stanford M. and William A. Douglass
1973 "Ethnicity: strategies of collective and individual impression management." Social Research 40:344–65.

Merton, Robert K.
1941 "Intermarriage and social structure: fact and theory." Psychiatry 4:361–74.
1957 Social Theory and Social Structure. Glencoe: The Free Press.

Miyamoto, S. Frank
1939 "Social solidarity among the Japanese in Seattle." University of Washington Publications in the Social Sciences 11:57–130.

Moore, Joan W.
1976 Mexican Americans. 2nd ed. Englewood Cliffs, New Jersey: Prentice-Hall.

Myers, Jerome K.
1949 "The differential time factor in assimilation." Doctoral dissertation, Department of Sociology, Yale University.
1950 "Assimilation to the ecological and social systems of a community." American Sociological Review 15:367–72.

Park, Robert Ezra
1914 "Racial assimilation in secondary groups." Publications of the American Sociological Society 8:66–72.
1950 Race and Culture. Glencoe: Free Press.

Park, Robert Ezra and Herbert A. Miller
1921 Old World Traits Transplanted. New York: Harper.

Petersen, William
1970 "Success story: Japanese American style." Pp. 169–78 in Minako Kurakawa (ed.), Minority Responses. New York: Random House.
1971 Japanese Americans: Oppression and Success. New York: Random House.

Rose, Jerry D.
1976 The Ethnic Dimension in Human Relations. Chicago: Rand McNally.

Sandberg, Neil C.
1974 Ethnic Identity and Assimilation: The Polish American Community. New York: Praeger.

Schneider, David M. and George C. Homans
1955 "Kinship terminology and the American kinship system." American Anthropologist 57:1194–208.

Spiro, Melford E.
1955 "The acculturation of American ethnic groups." American Anthropologist 57:1240–52.

Stuckert, Robert P.
1963 "Occupational mobility and family relationships." Social Forces 41:301–7.

Taft, Ronald
1957 "A psychological model for the study of social assimilation." Human Relations 10:141–56.

Tinker, John N.
1973 "Intermarriage and ethnic boundaries: the Japanese American case." Journal of Social Issues 29:49–66.

U.S. Bureau of the Census
1963 Nonwhite Population by Race. Final Report PC(2)-1C. Washington, D.C.: United States Government Printing Office.
1970 Marital Status. Subject Report PC-2-4C. Washington, D.C.: U.S. Government Printing Office.
1973a Characteristics of the Population. United States Summary, v. 1 pt. 1 sec. 2. Washington, D.C.: U.S. Government Printing Office.
1973b American Indians. Subject Report PC(2)-1F. Washington, D.C.: U.S. Government Printing Office.
1973c Negro Population. Subject Report PC(2)-1B. Washington, D.C.: U.S. Government Printing Office.
1973d Japanese, Chinese and Filipinos in the United States. Subject Report PC(2)-1G. Washington, D.C.: U.S. Government Printing Office.

Uyematsu, Amy
1971 "The emergence of yellow power in America." Pp. 9–13 in Amy Tachiki et al. (eds.), Roots: An Asian American Reader. Los Angeles: UCLA Asian American Studies Center.

Warner, W. Lloyd and Leo Srole
1945 The Social Systems of American Ethnic Groups. New Haven: Yale University Press.

Whyte, William Foote
1943 Street Corner Society. Chicago: University of Chicago Press.

Wittke, Carl
1956 The Irish in America. Baton Rouge: Louisiana State University Press.

Wong, Paul
1972 "The emergence of the Asian-American movement." Bridge 2:33–9.
1981 "The political culture and politics as a vocation: the Asian American case." In Young-Hwan Jo (ed.), Asian Americans and Politics, forthcoming.

A MULTILEVEL MODEL OF LIFE SATISFACTION: EFFECTS OF INDIVIDUAL CHARACTERISTICS AND NEIGHBORHOOD COMPOSITION*

ROBERTO M. FERNANDEZ
University of Chicago

JANE C. KULIK
Abt Associates

The effects of individual attributes, neighborhood context, and neighborhood social comparison on self-reported life satisfaction are examined for a sample of United States residents. We hypothesize that neighborhoods are important social contexts within which individuals draw satisfaction from life. In the individual-level model of life satisfaction, the results agree with past research that age, education, health, and marital status affect satisfaction. The effects of neighborhood context and social-comparison processes show that rural dwellers are more satisfied than city dwellers, while persons living in neighborhoods with a high cost of living are less satisfied. People whose incomes are below the neighborhood average may be less satisfied.

Many studies conclude that people assess the quality of their lives not in terms of their absolute level of material resources, but relative to social norms prescribing the level of resources they ought to have. Easterlin (1974:12) argues that consumption norms exist in societies as standards against which individuals measure their own achievements. As objective conditions change, so do social norms, thus explaining why more prosperous countries are no happier than poor ones. Duncan (1975) also finds that in addition to one's absolute level of income, satisfaction with standard of living is related to one's relative position in the income distribution. Similarly, Campbell, Converse, and Rodgers (1976:13–7) show that well-being is dependent on a person's comparison of his/her life conditions with internalized standards for his "most liked" experience.

These studies suggest that people use multiple reference standards to evaluate their positions. Such standards reveal "structural constraints" on attitudes and behavior, exerted by common values and status distributions (Blau, 1970:167–9; Merton, 1957:225–35). Yet few attempts have been made to examine simultaneously the effects of structural constraints, group structure, and individual attributes on well-being.

This paper explores the effects of individual attributes, neighborhood composition, and social comparison on self-reported life satisfaction. Urban sociologists (Suttles, 1972; Gans, 1962; Rainwater, 1970) have shown that neighborhoods function as social contexts that structure awareness of the world. Borrowing Easterlin's (1974) notion of consumption norms, we argue that individuals compare their material possessions and status to their neighbors' in evaluating how well off they are. We subscribe to Rainwater's (1974:23) claim that "consumption . . . locates the family in the stratification system;" and this location becomes most apparent to individuals within a neighborhood context.

We draw on models described by Firebaugh (1979) and Boyd and Iversen (1979:Chap. 3), and use regression

---

* Direct all correspondence to Roberto M. Fernandez, Department of Sociology, University of Chicago, Chicago, IL 60637.

Authorship is alphabetical. This work was done in conjunction with the Contextual Project at the National Opinion Research Center, and was funded by NSF Grant SOC 77-14515. Principal investigators are Arthur Stinchcombe and D. Garth Taylor. We would like to thank Chris Winship, Art Stinchcombe, François Nielsen, Lutz Erbring, and the anonymous *ASR* reviewers for their helpful suggestions and criticisms. Any remaining errors are ours alone.

methods to test propositions about the role of neighborhoods in generating standards of comparison by specifying a model which allows neighborhood characteristics to affect life satisfaction independent of individual attributes. Because individual and collective properties can interact in many ways (see Lazarsfeld and Menzel, 1961; Davis, Spaeth, and Huson, 1961), we construct explicit social-comparison indicators to measure an individual's distance from the neighborhood mean on income and age. The formal operationalization of this model is modified from Boyd and Iversen (1979:65–75) and Taylor (1973:1207–8).

Since misspecification can lead to mistaken contextual effects where a simpler, individual-level model would fit the data as well (Hauser, 1970, 1974), we first develop an individual-level model, using traditional variables that predict life satisfaction; we then add neighborhood-context variables and test for their independent effects.

We include family size among the individual-level regressors because past research has shown well-being is tempered by family circumstances (Wilkening and McGranahan, 1978:225, Table IX).[1] Although past research has not shown direct effects of gender on life satisfaction (e.g., Robinson and Shaver, 1973:17–9), we include sex among our predictors because a few studies have noted its conditioning effect on the relationship of other variables on well-being (Campbell, Converse, and Rodgers, 1976; Clemente and Sauer, 1976).

Previous studies have shown inconsistent results with regard to the influence of age on well-being (e.g., Phillips, 1967; Cantril, 1965:257–62; Clemente and Sauer, 1976:627–30; Campbell, Converse, and Rodgers, 1976:150–65; Alston and Dudley, 1973). These discrepancies probably result from the different instruments used and different populations under study (Alston and Dudley, 1973; Phillips, 1967). As Campbell and others have shown (Campbell, 1976; Campbell, Converse, and Rodgers, 1976:36–7), happiness and life satisfaction measures behave differently with respect to age: older persons are more satisfied but less happy. Gurin, Veroff, and Feld (1960:212–3 and 43, Table 2.9) also report older persons as more satisfied across a number of life areas but less happy.

The relationship between education and life satisfaction has been typically zero or slightly negative in past studies (Clemente and Sauer, 1976; Campbell, Converse, and Rodgers, 1976; Wilkening and McGranahan, 1978; but see Glenn and Weaver, 1981). With regard to race, whites have appeared more satisfied with their lives than nonwhites (Clemente and Sauer, 1976).

Health has displayed a strong, consistent positive relationship to well-being (Clemente and Sauer, 1976; Palmore and Luikart, 1972; Bradburn, 1969). Marital status has also surfaced as a reliable predictor of well-being. Clemente and Sauer (1976:627) found that married persons are more satisfied with their lives than those never married. Robinson and Shaver (1973:17–9) and Gurin, Veroff, and Feld (1960:230–8) found in addition that divorced or separated persons were less satisfied. Lastly, income has been shown to be positively related to a number of indicators of well-being (Easterlin, 1974; Bradburn, 1969; Cantril, 1965; Campbell, Converse, and Rodgers, 1976; Alston, Lowe, and Wrigley, 1975).

We investigate the effect of neighborhood context on life satisfaction by including the percentage of neighborhood residents who are white, the neighborhood's estimated cost of living, estimated income inequality within the neighborhood, and a dummy variable if the neighborhood is urban.

Percentage of whites may be a proxy for satisfaction-producing characteristics such as level of public services, school quality, or other factors which vary across neighborhoods but which are not measured explicitly in our survey. Further, we expect that neighborhoods with a high cost of living or with substantial income

---

[1] Rainwater's (1974:94–100) argument for considering the family as the consumption unit is the basis of our claim here. However, the larger body of microeconomic theory, particularly Becker's work on intrafamily allocation of time and interdependence of utilities, also provides impetus for these contentions (see Becker, 1977).

inequality might have less satisfied residents. The mechanism by which cost of living should operate is straightforward: persons living in high-cost neighborhoods have part of their material prosperity removed as a cost of residence. With regard to income inequality, we hypothesize that persons are more likely to be satisfied if they live in materially homogeneous surroundings, regardless of their own place in the neighborhood income distribution. Lastly, either because of self-selection of dissatisfied persons out of rural areas and the lower aspiration levels of those remaining, or because conditions in general are more favorable in rural areas, we expect city dwellers to be less satisfied than rural residents (see Wilkening and McGranahan, 1978:218).

In addition to the effects of neighborhood context per se, we are also interested in testing for the existence of neighborhood-level social-comparison processes and their effects on life satisfaction. Net of other variables in the model, individuals whose incomes are lower than the neighborhood average are hypothesized to feel "relatively deprived" and those whose incomes are above are "relatively gratified" (Davis, 1959, 1966). We can envision similar comparisons being made for age, where those above the mean are old relative to their neighborhoods and those below are young. It might be that individuals evaluate their satisfaction relative not only to their situation at an earlier point in time (c.f. Campbell, Converse, and Rodgers, 1976:177–81), but also to persons around them who might be in a different life-cycle stage. We hypothesize that persons older than the neighborhood average would receive a satisfaction bonus independent of their own age: they compare themselves to persons who have not yet undergone the life-stage transitions and feel an attendant sense of accomplishment and satisfaction.

DATA AND VARIABLES

The data for this study are taken from the National Opinion Research Center's Continuous National Survey (CNS). The survey was conducted between April 13, 1973 and May 30, 1974, with sampling broken into twelve cycles over that period. The sample ". . . was based on the NORC Master Probability Sample of Households—a multistage, stratified, full probability sample of all persons 18 years of age or older, living in households within the 48 contiguous United States" (NORC Report 125, 1974:i; see also Frankel and Kohnke, n.d., and King and Richards, 1972, for descriptions of the sampling frame).

This survey has two advantages for this analysis: the large case base (N = 7954) minimizes any problems due to sampling error in the dependent variable,[2] and the average number of persons sampled per segment of each Primary Sampling Unit is much larger (13.2) than in most surveys; this enables us to describe the distribution of the population within neighborhood contexts. At the same time, since NORC's sample selection procedure defines a segment as an area "consisting of approximately 100 households according to Census figures" (King and Richards, 1972:15), these units are small enough to serve as proxies for neighborhoods.[3] The CNS has 601 of such neighborhoods.

We use a self-reported measure of life satisfaction as our dependent variable. Responses to the question "Generally speaking, how satisfied would you say you are with your life as a whole—completely satisfied, very satisfied, moderately satisfied, slightly satisfied, or not

[2] Data for the CNS were collected in 12 cycles that were individually funded by different government agencies requesting information on various topics. Because of the specialized nature of some of the variables, many questions were not asked in all cycles. Restriction of cases to those with complete data for every variable used in the analysis leaves a working N of 5916.

[3] The alternative unit for which data were prepared by the Contextual Project at NORC is the Census tract, but two considerations make its use impractical. First, as each tract may include many thousands of households, we find it too large a unit for the examination of neighborhood-level context effects. Second, as only Standard Metropolitan Statistical Areas are included in Census tracts, all rural respondents are excluded. Since previous studies show that rural dwellers are more likely to be satisfied with their lives (Campbell, Converse, and Rodgers, 1976:236–7; Wilkening and McGranahan, 1978:219–20), we require a unit of aggregation that allows this finding to be tested.

at all satisfied?" were assigned integer values zero through four from least to most satisfied. While a larger number of response options would have been desirable, five categories have been shown to typically reveal over 90% of the total variance (Andrews and Withey, 1976:86; Doreian, 1972).

One problem that emerges in using self-reported measures of well-being is their reliability. Since we do not have a direct estimate of the reliability of our life-satisfaction item, we used the test-retest correlation for the life satisfaction item used in the Quality of American Life study (Campbell, Converse, and Rodgers, 1976). Though their question was worded somewhat differently, we believe it taps the same underlying construct as ours, overall life satisfaction. Using Campbell, Converse, and Rodgers' (1976) test-retest correlation of .43 over eight months, we see that the effect of instability and/or measurement error is quite sizeable. However, as the authors point out (p. 48), this is not atypical of well-being measures.

Because some change occurred over the eight months (Campbell, Converse, and Rodgers, 1976:48), the test-retest correlation should not be treated as a reliability coefficient (see Heise, 1969). But if we conceive of the underlying variable as a person's general disposition to express satisfaction with his life in view of his circumstances, and acknowledge that day-to-day fluctuations are not likely to be predicted with survey data, it becomes reasonable to treat the change over the eight month period together with the measurement error. Thus the figure of .43 should be a fair estimate of the reliability in the absence of alternatives.

We next infer an upper bound estimate of the change. We first derive an alternative lower bound estimate of the reliability for the single time point by regressing life satisfaction on a number of satisfaction and well-being items. These are the set of domain-satisfaction items—housing, neighborhood, work, financial, and marital satisfaction—(see Campbell, Converse, and Rodgers, 1976:61–2, for elaboration of the domain concept); the ten items comprising the Bradburn Affect Balance Scale; and the oft-used measure of self-reported happiness (Bradburn, 1969). As random measurement error cannot be predicted beyond chance (Stinchcombe, 1979), we take the multiple R of .54 between the life-satisfaction item and the set of well-being measures as an estimate of the reliability.

Since this estimate is based on data collected at a single time, the day-to-day fluctuations that affect life satisfaction also affect the other well-being measures and are reflected in the multiple R. Because we view the stable conception of life satisfaction as the more useful social indicator, it would be inappropriate to use this multiple R of .54 to correct the correlation matrix for attenuation due to measurement error. Though this figure is more conservative in that it would lead us to correct less (because .54 is larger than .43), it is based on a different conception of the underlying variable and hence is useless for that purpose. However, the measure does allow us to evaluate what portion of the error variance in our over-time measure of the reliability could reasonably be due to change.

Assuming the accuracy of the .54 figure, the error variance derived from that measure ($1 - .54^2 = .71$; see Walker and Lev, 1953:Chap. 12) can only be due to measurement error, since all the items were measured at the same time. If we assume that the three components of the variance in life satisfaction—true variance, variance due to measurement error, variance due to change over time—are uncorrelated, we can use the difference between the two estimates of the error variance to estimate the change component. The over-time estimate of the error variance, which includes both the variance due to error and change, is $1 - .43^2 = .82$. The difference, $.82 - .71 = .11$, can be treated as an estimate of the variance due to change in life satisfaction over time. This change includes both day-to-day fluctuations associated with our single-time idea of life satisfaction and the more drastic change implied in the more constant conception. If we assume conservatively that all of the change is drastic change, i.e., none of it is accounted for by day-to-day fluctuations, then we find that, at the maximum, change accounts for 13%

(.11/.82 = .134) of the total error variance observed over the eight-month period.

Since it is more likely that most of the variance due to change is attributable to day-to-day "noise" in the measure than to shifts in the more constant "signal," we adjust the correlations of the independent variables with the life-satisfaction measure above the diagonal in Table 1 for attenuation using the over-time reliability estimate (see Walker and Lev, 1953:299–306). The standard deviation presented above the diagonal was also corrected by removing the error variance (Walker and Lev, 1953:296–9). However, because of the tenuous nature of both reliability estimates and because correcting for attenuation in this case adds little new information, regressions are presented below using both corrected and uncorrected correlation matrices and standard deviations.[4]

Turning to the individual-level regressors, family size is measured as the number of people in the household. Gender is coded as a dummy variable (1 = male). Age is coded in number of years. The education variable is a continuous measure of "highest grade of school completed." Race is a dummy variable (1 = white). Hispanics are treated together with blacks, as preliminary analyses not reported here have shown that the results are unaffected by their omission. Health status is based on the question: "Compared to other people your age, would you say that your health is very good, good, fair, poor, or very poor?", coded 5 to 1. The wording is intended to separate the effects of age and health, and the modest correlation of −.23 between these two variables shown in Table 1 demonstrates this aim was at least partially achieved. As a further check on the joint effect of age and health a multiplicative interaction term (health × age) was included in preliminary analysis to determine if there was a satisfaction bonus attached to being a particularly healthy older person. Since the effect was insignificant when age and health were controlled, we have not included it here. Marital status is measured by two dummy variables, one for being married and the other for being divorced. The comparison group is those who have never married. Lastly, our income variable is total family income before taxes in 1972 for cycles 1–9 and 1973 for cycles 10–12. This was asked in categories ranging from under $1,000 to $25,000 and above. Each category was given the midpoint of the range, with $33,000 assigned to the highest category, to approximate a continuous measure. We use family income rather than individual income because it is a more accurate measure of the economic resources individuals have at hand. For example, the command that an unemployed family member has over monetary resources is more precisely rendered by total household income. Both because income is distributed lognormally in most survey data and because evidence suggests that the relationship between income and well-being is concave (Winship, 1976), the variable used here is transformed into its natural logarithm.

The neighborhood-level predictors—percentage of white residents, mean income, and mean age of residents—are constructed averages of within-neighborhood individual values. The two social-comparison indicators are taken as the difference between the individual's value on income and age and the neighborhood mean. The urban dummy was constructed by dichotomizing the interviewer's response to the query, "Would

[4] Correcting for attenuation in the dependent variable does not change the unstandardized coefficients, and since we are using the unadjusted standard errors to evaluate the significance of the coefficients (see Walker and Lev, 1953:306–7), the correction procedure cannot introduce a significant effect that was not present in the original data. Correcting for attenuation in the dependent variable does change the standardized coefficients. But since each coefficient is adjusted by a constant—the adjusted standard deviation of Y—there is no change in the relative magnitudes of the effects, but only in their absolute sizes. Under the assumption that all measurement error has been removed from the dependent variable, we get more accurate estimates of the standardized coefficients, though assessment of statistical significance is still a function of the uncorrected standard errors. Aside from better estimation of the standardized coefficients, the only other reason to correct for attenuation is to see how well the model fits the data after removing the measurement error. While it is more encouraging to note the larger R^2 that the corrected model yields, our reliability measure is provisional and must affect confidence in the R^2.

Table 1. Correlations, Means and Standard Deviations for the Multilevel Model (Coefficients above the diagonal have been corrected for attenuation; those below are uncorrected)

	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	Mean	SD
Y: Life satisfaction		.005	.020	.096	−.016	.077	−.104	.301	.216	−.194	.106	.115	−.138	.091	.091	−.006	2.841	.583
X1: Family size	.003		.023	−.378	.041	−.130	−.048	.092	.318	−.326	.277	.264	.001	−.341	−.081	−.049		
X2: Sex = Male	.013			−.028	.062	.033	−.025	.051	.110	−.202	.137	.137	.014	−.023	.042	−.009		
X3: Age	.063				−.341	.080	−.038	−.234	−.100	.344	−.259	−.253	−.019	.917	.046	.071		
X4: Education	−.011					.204	.156	.279	.072	−.175	.442	.291	.140	−.317	.206	−.124		
X5: Race = White	.050						−.097	.097	.080	−.067	.218	.049	.017	.054	.771	−.045		
X6: Urban	−.068							.070	−.097	.056	.080	.011	.290	−.010	−.126	−.176		
X7: Health	.196								.077	−.124	.271	.200	.063	−.212	.105	−.070		
X8: Married	.140									−.713	.351	.308	−.062	−.073	.095	−.080		
X9: Divorced/separated	−.126										−.369	−.333	.013	.313	−.095	.078		
X10: Ln Family income	.069											.824	.217	−.219	.232	−.248		
X11: D1[a]	.075												−.001	−.275	.0001	.001		
X12: Cost of living	−.089													.015	.025	−.228		
X13: D2[b]	.059														.009	−.006		
X14: Percent white	.059															−.066		
X15: Inequality ratio	−.004																	
Mean	2.841	3.059	.431	43.583	11.604	.865	.757	4.169	.684	.190	9.057	−.014	9.411	.002	.862	.617		
SD	.889	1.726	.495	17.471	3.265	.342	.429	.909	.465	.393	.893	.736	.088	15.999	.266	.210		
N	5916																	

[a] D1 = difference between own income and neighborhood mean income.
[b] D2 = difference between own age and neighborhood mean age.

you consider the place where R lives to be a city, a suburb, a town, a rural area, or a farm?"; the first three choices are considered urban. Preliminary analysis has shown that the adjusted means for the places comprising the urban category are sufficiently equal to justify this aggregation. The cost of living measure is derived from published tables in the 1973 and 1974 Statistical Abstract, which present average cost of living for a family of four in approximately 50 cities and rural counties. Neighborhoods were assigned the logarithm of the cost-of-living estimate by matching their Primary Sampling Unit codes with the appropriate city or county figure. Finally, the income inequality measure is a variant of Atkinson's measure (see Atkinson, 1975:45–9), roughly interpreted as the ratio of the geometric to the arithmetic mean of the income distribution in each neighborhood. This ratio varies between zero and one, where one signifies complete equality of income.[5]

ANALYSIS

Self-reported health status is the best predictor of life satisfaction (see Tables 1 and 2); healthier persons are significantly more satisfied with life. Comparison of the correlation coefficient with the uncorrected standardized coefficient shows that the direct effect is a bit larger. The relationship between health and life satisfaction is apparently suppressed by the effects of the other variables in the model, most notably age (note the correlation of −.234 between health and age). When comparing persons of the same age, the relationship between health and life satisfaction is even stronger than in the population at large. The second best predictor of life satisfaction is age: older persons are significantly more satisfied than younger, perhaps because of a reevaluation of expectations that accompanies this stage of the life cycle (see Campbell, Converse, and Rodgers, 1976).

Table 2. Determinants of Life Satisfaction: Individual-Level Model (Unstandardized and standardized regression coefficients, and standardized coefficients corrected for attenuation. Standard errors in parentheses.)

Independent Variable	Metric Coefficients	Standardized Coefficients: Uncorrected	Standardized Coefficients: Corrected
Family size	−.012 (.008)	−.023	−.036
Sex = Male	−.041 (.023)	−.023	−.036
Age	.007* (.0008)	.135	.207
Education	−.015* (.004)	−.056	−.086
Race = White	.031 (.035)	.012	.018
Health status	.218* (.013)	.223	.343
Married	.119* (.036)	.062	.096
Divorced/separated	−.267* (.045)	−.118	−.181
Ln of family income	.010 (.016)	.010	.015
Intercept	1.720		
R^2		.077	.182

* $p < .05$ (as derived from the standard errors).

Our findings agree with those of Wilkening and McGranahan (1978) with respect to the sign of the coefficient for family size. Persons in larger families are less satisfied with their lives. However, while the authors find family size significant in predicting 4 out of 5 of their life satisfaction items (pp. 225–9), the slope in our model is insignificant.

We find a significant negative effect of education on satisfaction. The usual interpretation of this result is that education raises aspirations that are not easily fulfilled, so that higher education produces dissatisfaction with life (see Campbell, Converse, and Rodgers, 1976:172–99).[6] In

[5] The value of E, the weight given to the bottom of the income distribution in the calculation of the income inequality measure was set at −.25. Varying E up to +2 does not change the results reported here. For a detailed discussion of Atkinson's measure and a comparison of it with other measures of inequality, see Schwartz and Winship (1980).

[6] Following this interpretation, we hypothesized that highly educated nonwhites would exhibit the highest degree of dissatisfaction because past constraints on nonwhite achievement make a high level of educational attainment an aspiration-swelling experience for this group. We tested this hypothesis in preliminary analysis by entering an interaction term coded 1 if highly educated (13+ years) and nonwhite,

accord with past research, married persons are more satisfied and divorced or separated persons less satisfied than those who have never married.

The magnitude of the sex difference in life satisfaction is both substantively and statistically insignificant. The partial effect of race on satisfaction also fails significance at the usual (.05) level. Analyses not reported here show that this remains the case when Hispanics are removed from the sample. Lastly, our measure of family income is insignificant. Though the sign of the effect is as expected, its small size is somewhat of an anomaly in light of previous investigations. Because sex, race, and income are correlated with our contextual variables (see Table 1), we will continue to control for them despite their statistical insignificance in the individual-level model.

When we examine correlations between the individual-level and neighborhood-context variables, several relationships stand out (Table 3).[7] The .771 correlation between race and neighborhood racial composition indicates that both blacks and whites tend to live in racially homogeneous surroundings. Also, exceptionally high correlations between age and income and the respective social-comparison variables reveal that most of the variance in both age and income occurs within neighborhoods. Because neighborhoods do not vary much in their means for age and income, it might be difficult to separate the effects of absolute from relative levels of age and income. The high degree of segregation might also hamper our assessment of the relative effects of the individual race and the neighborhood racial-composition indicator.

Table 3. Determinants of Life Satisfaction: Multi-Level Model (Unstandardized and standardized regression coefficients, and standardized coefficients corrected for attenuation. Standard errors in parentheses.)

Independent Variable	Metric Coefficients	Standardized Coefficients: Uncorrected	Standardized Coefficients: Corrected
Family size	−.013 (.007)	−.025	−.039
Sex = Male	−.046* (.023)	−.026	−.039
Age	.006* (.002)	.113	.173
Education	−.011* (.004)	−.042	−.064
Race = White	−.045 (.052)	−.017	−.026
Health status	.220* (.013)	.225	.346
Married	.083* (.036)	.043	.067
Divorced/separated	−.269* (.045)	−.119	−.182
Ln family income	−.015 (.032)	−.015	−.024
Urban	−.082* (.028)	−.040	−.061
D_1	.061 (.034)	.051	.078
Cost of living	−.816* (.143)	−.081	−.125
Percent white residents	.116 (.070)	.035	.053
D_2	.002 (.002)	.033	.050
Inequality ratio	−.074 (.060)	−.017	−.027
Intercept	9.727		
R^2		.088	.208

* $p < .05$ (as derived from standard errors).

NOTE: D_1 = difference between own income and mean income of the neighborhood.

D_2 = difference between own age and mean age of neighborhood.

As anticipated, persons living in neighborhoods with a high cost of living are less satisfied with their lives than those in low-cost neighborhoods. The straightforward economic mechanism described above would account for this dissatisfaction: holding family income constant, persons living in high-cost neighborhoods are less affluent than those in low-cost areas. Persons living in cities are significantly less satisfied than those living in rural areas (standardized direct effect −.040).

Though failing significance at conventional levels, the coefficient of .061 on the income-difference term suggests that persons whose incomes are above the neighborhood average are "relatively gratified" and exhibit high levels of satis-

and 0 otherwise. The effect was solidly insignificant and created a serious multicollinearity problem with the race variable because of the small number of nonwhites who were highly educated, so the interaction term was dropped from further analyses.

[7] Since all correlations were computed at the individual level, each neighborhood is in effect weighted by its number of respondents.

faction, while persons with incomes below the mean are "relatively deprived" and therefore less satisfied. Contrary to our expectations, the close-to-zero coefficient on the age-difference term indicates that persons older than the neighborhood average are no more satisfied than those who are younger.

There is a slight tendency for persons residing in predominantly white neighborhoods to be more satisfied (.116), though this is not a statistically reliable effect (standard error .070). The percentage of whites might be tapping satisfaction-producing neighborhood characteristics, such as level of services, for which we do not have explicit measures. Lastly, as the small and insignificant coefficient shows, persons in neighborhoods with more equal income distributions are not more satisfied.

Adding these variables as controls also produces changes in some of the slopes of the individual-level variables. In particular, the slight tendency in the individual-level model for women to be more satisfied becomes statistically significant in the contextual model by way of a change in the third decimal place. Especially as gender is virtually uncorrelated with the neighborhood variables, we have no substantive explanation for this.

Though statistically significant in both models, married persons appear a bit more satisfied in the individual-level model. Married persons are more likely to be above the neighborhood mean on family income ($r = .308$), probably because larger families require more income and have more earners (note the correlation of .277 between family size and family income). The introduction of the social-comparison indicator for income, then, explains in part why married persons are more satisfied: persons with families are likely to be relatively gratified because their income is higher than the mean neighborhood income.

In the individual-level model, whites are slightly more satisfied than nonwhites; in the multilevel model, nonwhites show a slight tendency to report themselves as satisfied once the effects of all other variables are partialled out. Also, while wealthier respondents are found to be less satisfied in our contextual model, the opposite is true in our individual-level model. However, since these anomalies are probably due to the high degree of collinearity between the individual and neighborhood race and income indicators,[8] and the race and income effects are statistically insignificant in both the individual and multilevel models, we interpret these effects as nil in both models.

The remaining individual-level effects are essentially unchanged after the introduction of the neighborhood contextual variables, indicating that, aside from the anomalies just noted, the neighborhood characteristics are adding new information to the regressions explaining life satisfaction. This becomes clearer when we consider the R^2s in Tables 2 and 3. Although higher absolute R^2s would have been more satisfying, F tests show that the variance explained in life satisfaction for both corrected and uncorrected models at each level of analysis is highly significant. More importantly, for both the corrected and uncorrected models, the increment in variance explained by the addition of the neighborhood characteristics and social-comparison indicators is statistically significant (F tests of 32.28 and 11.86, each with 6 and 5900 degrees of freedom). This indicates that the neighborhood contextual variables are tapping sources of life satisfaction that are independent of the attributes in our individual-level model.

SUMMARY AND CONCLUSION

In this paper, we tested a number of hypotheses regarding the effects of neighborhood characteristics and social-comparison processes that we reasoned should operate on respondents' self-reported life satisfaction. Borrowing Easterlin's (1974) notion of consumption norms, and building on the work of urban sociologists (Suttles, 1972; Gans, 1962; Rainwater, 1970), we have argued that neighborhoods are social contexts within

[8] Analyses not reported here lend support to this inference. In both cases, inclusion of the contextual variable swells the standard error of the individual-level coefficients considerably.

which individuals draw satisfaction from life. We hypothesized that various characteristics of neighborhoods and individuals' social comparisons to neighborhood-level standards should affect respondents' life satisfaction.

To test these propositions, we began by specifying an individual-level model of life satisfaction by drawing on past research conducted at the individual-level. In large part, our results agree with extant research and show effects of age, education, health, and marital status on satisfaction.

We next examined the effects of neighborhood context and social comparison processes. We found rural dwellers to be more satisfied than city dwellers, and respondents living in neighborhoods with a high cost of living to be less satisfied. In addition, controlling for other individual and neighborhood attributes, we noted a small and perhaps unreliable effect of relative deprivation, whereby individuals with incomes below the neighborhood average appear less satisfied. The coefficients for the remaining neighborhood-level variables were found to be insignificant. Lastly, the introduction of the neighborhood indicators left the statistically reliable effects of the individual model largely unchanged, indicating that the above neighborhood variables are tapping independent sources of satisfaction.

The apparent insignificance of the social-comparison indicators and the absence of composition effects for race and income inequality are not disturbing. Our neighborhood-composition indicators describe average measurable attributes of the neighborhood and as such are fairly objective. Yet these appear not to be the qualities of the neighborhood which are most salient to individuals in producing satisfaction with life. Other social structural variables such as the degree of social interaction and number of friendship and kinship ties may prove to be aspects of the neighborhood context which have the most significant impact on individuals' self-reported life satisfaction.

Also, social-comparison processes may be operating at different levels which crosscut or suppress the effects of comparisons made within neighborhoods. Contextual variables at the work-group or friendship-group level are the most likely to qualify the findings we report here. Though our data could not test these alternative hypotheses, we have shown evidence of some context effects at the neighborhood-level on life satisfaction.

REFERENCES

Alston, Jon and Charles Dudley
1973 "Age, occupation, and life satisfaction." The Gerontologist 13:58–61.

Alston, Jon, George Lowe, and Alice Wrigley
1974 "Socioeconomic correlates for four dimensions of self-perceived satisfaction, 1972." Human Organization 33:99–102.

Andrews, Frank and Stephen Withey
1976 Social Indicators of Well-Being: American Perceptions of Life Quality. New York: Plenum.

Atkinson, A. B.
1975 The Economics of Inequality. Oxford: Oxford University Press.

Becker, Gary
1977 The Economic Approach to Human Behavior. Chicago: University of Chicago Press.

Blau, Peter M.
1960 "Structural effects." American Sociological Review 25:178–93.

Boyd, Lawrence and Gudmund Iversen
1979 Contextual Analysis: Concepts and Statistical Techniques. Belmont, California: Wadsworth.

Bradburn, Norman
1969 The Structure of Psychological Well-Being. Chicago: Aldine.

Campbell, Angus
1976 "Subjective measures of well-being." American Psychologist 31:117–24.

Campbell, Angus, Philip Converse, and Willard Rodgers
1976 The Quality of American Life. New York: Russell Sage Foundation.

Cantril, Hadley
1965 The Pattern of Human Concerns. New Brunswick, New Jersey: Rutgers University Press.

Clemente, Frank and William Sauer
1976 "Life satisfaction in the United States." Social Forces 54:621–31.

Davis, James A.
1959 "A formal interpretation of the theory of relative deprivation." Sociometry 22:280–96.
1966 "The campus as a frog pond: an application of the theory of relative deprivation to career decisions of college men." American Journal of Sociology 72:17–31.

Davis, James A., Joe Spaeth, and Carolyn Huson
1961 "A technique for analyzing the effects of group composition." American Sociological Review 26:215–25.

Doreian, Patrick
1972 "Multivariate analysis of categorical data." Quality and Quantity 6:253–72.

Duncan, Otis Dudley
1975 "Does money buy satisfaction?" Social Indicators Research 2:267–74.
Easterlin, Richard
1974 "Does economic growth improve the human lot? Some empirical evidence." Pp. 89–125 in Paul A. David and Melvin Reder (eds.), Nations and Households in Economic Growth. New York: Academic.
Firebaugh, Glenn
1979 "Assessing group effects: a comparison of two methods." Sociological Methods and Research 7:384–95.
Frankel, Martin and Luane Kohnke
n.d. "NORC national probability sample frame." Unpublished. Chicago: National Opinion Research Center.
Gans, Herbert
1962 The Urban Villagers. New York: Free Press.
Glenn, Norval and Charles Weaver
1981 "Education's effects on psychological well-being." Public Opinion Quarterly 45:22–39.
Gurin, Gerald, Joseph Veroff, and S. Feld
1960 Americans View Their Mental Health. New York: Basic.
Hauser, Robert M.
1970 "Context and consex: a cautionary tale." American Journal of Sociology 75:645–64.
1974 "Contextual analysis revisited." Sociological Methods and Research 2:365–75.
Heise, David
1969 "Separating reliability and stability in test-retest correlations." American Sociological Review 34:93–101.
King, Benjamin and Carol Richards
1972 "The 1972 NORC national probability sample." Unpublished. Chicago: National Opinion Research Center.
Lazarsfeld, Paul and Herbert Menzel
1961 "On the relation between individual and collective properties." Pp. 422–40 in Amitai Etzioni (ed.) Complex Organizations: A Sociological Reader. New York: Holt, Rhinehart, and Winston.
Merton, Robert K.
1957 Social Theory and Social Structure. New York: Free Press.
NORC Report Number 125.
1974 Continuous National Survey: A Comparison of Questionnaire Items, Cycles 1 Through 12. Chicago: National Opinion Research Center.
Palmore, E. and C. Luikart
1972 "Health and social factors related to life satisfaction." Journal of Health and Social Behavior 13:68–80.
Phillips, Derek C.
1967 "Social participation and happiness." American Journal of Sociology 72:479–88.
Rainwater, Lee
1970 Behind Ghetto Walls: Black Families in a Federal Slum. Chicago: Aldine.
1974 What Money Buys. New York: Basic.
Robinson, John and Phillip R. Shaver
1973 Measures of Social Psychological Attitudes. Ann Arbor: Institute for Social Research, University of Michigan.
Schwartz, Joseph and Christopher Winship
1980 "The welfare approach to measuring inequality." Pp. 1–36 in Samuel Leinhardt (ed.) Sociological Methodology. San Francisco: Jossey-Bass.
Stinchcombe, Arthur
1979 Personal Communication.
Suttles, Gerald D.
1972 The Social Construction of Communities. Chicago: University of Chicago Press.
Taylor, Howard F.
1973 "Linear models of consistency: some extensions of Blalock's strategy." American Journal of Sociology 78:1192–225.
United States Bureau of Census
1973 Statistical Abstract of the United States, 1973. Washington, D.C.: U.S. Bureau of Census.
1974 Statistical Abstract of the United States, 1974. Washington, D.C.: U.S. Bureau of Census.
Walker, Helen and Joseph Lev
1953 Statistical Inference. New York: Holt, Rhinehart and Winston.
Wilkening, E. A. and David McGranahan
1978 "Correlates of subjective well-being in Northern Wisconsin." Social Indicators Research 5:211–34.
Winship, Christopher
1976 "Psychological well-being, income and income inequality." Unpublished.

INTERLOCKING DIRECTORATES AND INTEREST GROUP FORMATION*

BETH MINTZ
University of Vermont

MICHAEL SCHWARTZ
State University of New York, Stony Brook

This paper uses data on interlocking directorates to test three theories of corporate organization: managerialism, coalition theory, and the theory of finance capital. Findings suggest that the modern corporation is not an autonomous unit as suggested by managerialism, that firms do not form flexible alliances which pursue mutual interests as implied by coalition theory, and that the interest groups of traditional finance capital theory do not characterize the interlock network of the 1960s. Instead, the system is dominated by a handful of interconnected major New York commercial banks and insurance companies which form the center of an integrated national network.

Since World War II, social scientists have engaged in an evolving debate over the structure of power in the United States. Beginning with disputes over the locus of community decision making (e.g. Hunter, 1953 versus Dahl, 1961), the controversy shifted to the democratic responsivenes of the national political elite (e.g. Mills, 1956 versus Truman, 1951), and then to the nature of government connections to large corporations (e.g. Domhoff, 1970 versus Rose, 1967). More recently, research has focused on intercorporate relations, since the nature and degree of business unity in large part determines the amount of political leverage available to corporate leadership (Koenig and Gogel, 1981; Mintz and Schwartz, forthcoming; Moore, 1979; Ratcliff, 1979–1980; Ratcliff, et al., 1979; Useem, 1979).

The present study follows this trend. By analyzing patterns of interlocking directorates, we evaluate the structure of relationships among major American corporations, assess the degree of interest-group formation within the business community, and use these results to assess contending theories of intercorporate coordination and competition.

INTERLOCKING DIRECTORATES AS INDICATORS OF THE STRUCTURE OF INTERCORPORATE INTERACTION

Interlocking directorates have been used in many contexts to evaluate corporate interaction patterns. In 1913, the Pujo Committee used interlocks to assess the role of financial institutions in the American economy. This investigation, coupled with other, nongovernmental research, contributed to the proscription of certain types of director interlocks in the Clayton Antitrust Act (Brandeis, 1914; Bunting, 1979; Kotz, 1978; Mizruchi, 1980, 1982).

In the late 1930s, another federally sponsored interlock investigation produced Paul Sweezy's influential application of finance-capital theory to the American economy (N.R.C., 1939). Since then, government research has regularly used director interchanges as measures of intercorporate coordination (Federal Trade Commission, 1951; U.S. Congress, 1965, 1968).

---

* Direct all correspondence to Beth Mintz, Department of Sociology, University of Vermont, 31 South Prospect Street, Burlington, VT 05405.

This research was supported by the National Science Foundation Grant Program under Award No. FOC 73-05606 and by the John Simon Guggenheim Memorial Foundation. The authors would like to thank William Atwood, James Bearden, Dan Clawson, Peter Freitag, Carol Hendricks, Christine Thornton and Eugene Weinstein. They would also like to thank the anonymous reviewers for helpful comments.

Until the 1960s, Marxists were the only social scientists who considered interlocks useful diagnostic tools for intercorporate research (Rochester, 1936; Perlo, 1957). However, since Warner and Unwalla (1967) demonstrated the pervasiveness of director exchanges and Dooley (1969) documented their stability over time, a great many researchers—non-Marxist and Marxist alike—have entered the field and refined both the technology and the theory of director interties.[1]

This blossoming of interlock research has laid to rest the longstanding critique, associated most directly with Myles Mace (1971), that interlocks have little significance in the corporate world. The elegant structures uncovered by Levine (1972) and others; the correlations between interlocks and a wide range of corporate activities, including dividend payout rates and short term indebtedness (Pfeffer and Salancik, 1978; Norich, 1980; Pennings, 1980; Gogel and Koenig, 1981); and the detailed exploration of corporate interaction patterns by European researchers (Mokken and Stokman, 1979; Fennema, forthcoming; Scott and Hughes, 1976) have demonstrated the usefulness of interlock studies in analyzing business structure.

These results have led many researchers to assume or assert that most interlocks represent dyadic "traces of power": one firm is wholly or partially controlled by the other, the two intertied firms are jointly controlled by a third party, or the two companies have at least a quasi-permanent structural alliance.[2] This assumption, however, has been challenged by recent research. Palmer's 1980 study of the meaning of interlocks found that less than 25% of all director exchanges represent dyadic traces of power. The remaining ties reflect a combination of social networks, family ties, local and national elite groupings, interpersonal economic alliances, and cross-sectoral economic interdependencies. Thus, two firms may be interlocked because they are actively collaborating, because their top leadership is interpersonally connected, because the two firms are in complementary industries which benefit from detailed interchanges of information, or from some combination of the above.

This conclusion is reinforced by many other studies of interlock networks. Pfeffer and Salancik (1978), Gogel (1976) and others have demonstrated that structural interdependence between economic sectors correlates with intersectoral interlocking (even when the specific firms do not collaborate). Mizruchi (1980, 1982) has shown that the changing roles of railroads, investment banks, and other types of firms are correlated with changes in interlock position, even in the absence of control relationships. Mintz and Schwartz (forthcoming) have demonstrated that bank centrality in corporate networks is a function of bank domination of capital flows and not a reflection of specific control relations.

These results suggest that interlocks are most usefully analyzed at the aggregate level. A great number of director exchanges implies significant intercorporate influence, significant economic interdependency, or an intermeshed social network of corporate leaders. Any combination of these underlying realities implies a high potential for common action. A specific firm will be heavily interlocked if it has a great many structural ties to other firms, if it is the center of a group of economic interdependencies, or if its leadership are personally connected to a great many other leaders. In any of these situations, highly interlocked companies may be central to coordinated action. Thus, interlock patterns are useful diagnostics for the existence of ongoing and potential

[1] See, for example, Aldrich and Pfeffer (1976), Allen (1974, 1978), Bearden et al. (1975), Berkowitz et al. (1979), Bunting (1976–77), Bunting and Barbour (1971), Burt (1979, 1980), Levine (1972, 1977), Mariolis (1975, 1978), Mintz and Schwartz (1981), Mizruchi (1980), Mizruchi and Bunting (1979), Mokken and Stokman (1974b, 1979), Palmer (1980a, 1980b), Pennings (1980), and Sonquist and Koenig (1975). For reviews of interlock research see Fennema and Schijf (1979), Scott (1979), and Soref (1979).

[2] The term "traces of power" was coined by Mokken and Stokman (1974a). Their assumption is better grounded than most American research, since they offer substantial evidence that interlock networks in the Netherlands do represent specific structural coordination.

relationships among major corporations. Patterns of intercorporate interaction are buried in the intertie system, and specially designed methods make these traces recoverable.

CORPORATE INTERLOCKS AND MANAGERIALISM

Differing theories of the firm contain different assumptions about the nature of interorganizational relationships and imply different models of corporate power. Managerialism, the dominant paradigm until the 1970s, posits interfirm autonomy based on the independence of top management from control by stockholders and lenders (Berle and Means, 1932; Kaysen, 1957; Larner, 1970; Galbraith, 1967; Bell, 1961). Companies, especially the very largest, become self-sufficient and self-reliant.[1] They are no longer subservient to outside finance or ownership interest and do not require alliances for maximizing profit or fending off aggressive competition.

Managerialism argues that the modern firm has little incentive to create interorganizational ties and is unlikely to be dramatically influenced by outside forces. It implies a corporate structure characterized by fragmentation—a collection of basically independent units pursuing individual interests—and, therefore, finds little basis for coalition formation. On a societal level, this implies that corporate power is dispersed and the danger of corporate domination is minimal.

To test the notion of corporate autonomy—to determine if the network of interlocking directorates among major American companies reflects an underlying system of distinct corporate actors free from significant interdependencies—we evaluate the pervasiveness and distribution of interlocks within the network. A highly integrated system of director exchanges is counterevidence to the managerialist image of the isolated, independent corporation.

However, even if a dense intertie network is found, this arrangement could still support the managerial view of declining unity and increasing independence of major firms, since a dense network might reflect a concentration of institutional connections among less independent, smaller sized companies. The smaller firms in our population might be tied together by traditional sorts of interdependencies with lenders, business partners, and outside representatives; while the larger corporations, with dispersed stock and management control, would have fewer interfirm relations and, therefore, fewer substantive interlocks. These larger corporations could be expected to retain some ties to smaller companies, while remaining less involved in the system as a whole. Since all managerialists argue that stock dispersal—and therefore managerial control and corporate independence—is correlated with company size (Berle and Means, 1932; Larner, 1970), managerialism hypothesizes that the larger corporations in our population should have fewer interlocks than smaller firms.

The Data

To test managerialism and the models discussed later in this paper, we use data on interlocking directorates among the largest companies in the United States of the 1960s, as listed by *Fortune* magazine. The data set includes the largest publicly owned industrial firms, commercial banks, insurance companies, transportation firms, utilities, merchandisers, diversified financial institutions, and nonclassifiable "miscellaneous" companies. An additional set of investment banks supplemented the *Fortune* listings. Information on board membership was compiled for the years 1962 and 1966 and taken from *Standard and Poor's Register of Corporations, Directors and Executives; Dun and Bradstreet's Million Dollar Directory;* the Moody's Manuals, *Directory of Directors for New York,* and Securities and Exchange Commission records.

The number of companies under consideration changed from year to year, as certain firms were acquired, merged, or disappeared. In the 1962 compilation, 1,131 different companies were included, while the 1966 data contained 1,111 firms. The number of exact positions also varied from year to year. In 1962, 13,574 individuals occupied 15,073 positions including

Table 1. Correlation Between Corporate Size and Number of Interlocks, 1962

Type of Corporation	# of Interlocks N	Mean	SD	Correlation Between Size and # of Interlocks[a]
Industrial	497	8.8	8.974	.457
Life insurance	50	13.9	16.244	.536
Commercial bank	49	25.9	19.823	.579
Utility	49	10.9	9.415	.481
Transportation	47	11.3	11.521	.670
Merchandising	39	8.4	8.668	.346

NOTE: Size is measured by rank in Fortune 500 listing for each type. Thus, the number-one corporation is the largest, number two is second, etc. The correlation is performed on industry rank, rather than assets or other ranking criteria, because of skewness in actual size distributions. Use of rank yields correlations of approximately the same magnitude as those between number of interlocks and log of size.

[a] Positive correlations indicate that larger firms have more interlocks. Technically, correlations between rank and number of interlocks would produce negative correlations if larger firms had more interlocks, since the "lower" rank numbers—1, 2, 3, etc.—would be associated with more interlocks. We report these as positive "size" correlations instead of negative "rank" correlations because it avoids the confusion of "reversed intuition."

5,699 interlocks among 1,003 different corporations. For a complete description of the data set, see Bearden et al. (1975) and Mintz (1978).

Results

The network of interlocking directorates gives little support to the idea of autonomous independent units. In 1962, 1,003 of the 1,131 companies under investigation (88.6%) maintained at least one intertie, while the total number of links, 5,699, represented an average of 10.0 per firm. These connections were not small clusters of companies linked together and isolated from other groupings: 989 of the corporations formed a single continuous system.

Moreover, as earlier research has suggested, the largest companies were the most intertied (Dooley, 1969; Allen, 1974; Mariolis, 1975). Table 1 presents the correlation between corporate rank in the 1962 *Fortune* listing and the number of interlocks maintained in that year. In every category, there is a moderate to strong positive correlation between company size and interlock frequency. The larger organizations do not extricate themselves from the network. Quite the contrary: as Table 2 demonstrates, the five largest firms in each category are far more interlocked than the industry average. Further, this is not a case of a few major corporations tying repeatedly into smaller enterprises. As we document later, larger firms are tied to each other in a tight, well-integrated network.

Approaching the problem from another vantage point, we find that the 128 isolates in the system are almost all smaller, less important companies; corporations which grow rapidly develop more interlocks and corporations which shrink typically lose ties. No matter how the problem is con-

Table 2. Interlock Frequency Average by Type of Corporation, 1962

Type of Corporation	# of Firms Per Type	Average # of Ties Per Corporation Type X̄	S.D.	Average # of Ties for the 5 Largest Firms in Each Type X̄	S.D.
Industrial	689	8.8	8.974	22.8	14.132
Transportation	73	11.3	11.521	26.8	11.945
Commercial bank	66	25.9	19.823	47.0	13.674
Merchandising	65	8.4	8.602	15.0	10.440
Utility	59	10.9	9.415	23.2	12.377
Life insurance	55	13.9	16.224	39.4	18.968

sidered, the larger firms are found to be more interconnected than the smaller units. This result strongly suggests that larger companies are more, not less, involved in intercorporate relationships and that the managerialist portrait of splendid isolation is simply not consistent with the evidence.

CORPORATE INTERLOCKS AND COALITION THEORY

A second model of intercorporate interaction argues that, despite managerial autonomy and conflicts of interest, the giant American company generally recognizes its dependence on—and its need for joint action with—other firms. As Williamson (1965:580) suggests, "oligopolists can be assumed to be continuously aware of their interdependency relationship and of the collective advantage of pursuing a qualified joint profit maximizing strategy . . ." Resource-dependency theory applies a similar logic. Aldrich (1979:267) notes that "one consequence of competition and sharing of scarce resources is the development of dependencies of some organizations on others." Interorganizational relationships are established to ensure a supply of these scarce resources (see Aldrich and Pfeffer, 1976; Pfeffer and Salancik, 1978).

These arguments imply a second model of intercorporate interaction which assumes the formation of alliances along industry or market lines. For convenience we call this "coalition theory," a view which asserts that firms with overlapping economic interests tend to interconnect, facilitating the pursuit of shared goals. Corporate structure should be characterized by coalition formation and at least temporary alliances among companies with shared interests.

The implications of this model are more ominous for political democracy than those derived from managerialism. When whole industries unite in pursuit of collective interests, their potential for successful intervention in government is greatly increased. Auto company coalitions, for example, have proven to be formidable political adversaries, both in their short term delay of pollution control and in their successful long term opposition to public rail transportation (Whitt, 1976, 1982; Yago, 1980). Such alliances, however, may be unstable and short lived. As conditions change, coalitions disaggregate; the fates of individual firms diverge and alliances change to reflect current needs. Thus, although the power base of the coalition is potentially vast, its effects on the larger society may be limited by the temporary nature of interest overlaps.

To test coalition theory, we search for groups of companies organized into functional clusters along market lines. The model suggests that the most powerful companies of an industry or sector maintain connections to a multitude of less important firms. General Motors, for example, might occupy a position at the center of a group of companies directly related to the auto industry (see Figure 1). Satellites might include: Libbey-Owens-Ford, the leading domestic windshield-glass producer; Firestone, Uniroyal, or some other major tire manufacturer; North American Rockwell, an automobile components producer; and National Steel, a major auto supplier. One might also find other arrangements, overlaid on top of these industrial groupings. Connections based on reciprocity might tie General Motors to Equitable Life, a leading source of funding, and to U.S. Steel, a representative of an industry with which auto has many common interests. Each of these com-

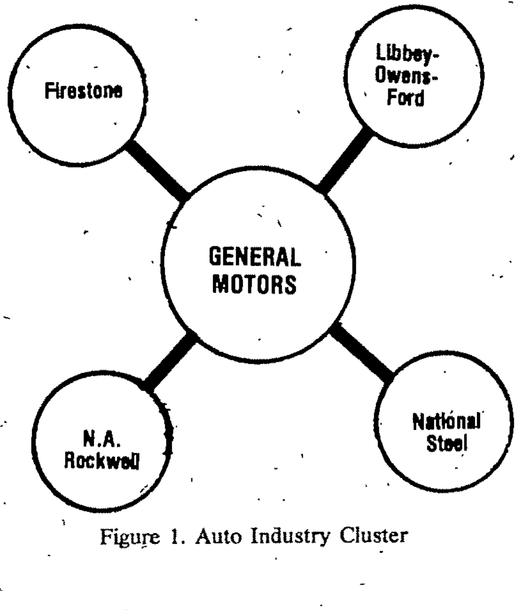


Figure 1. Auto Industry Cluster

panies could be the center of a group of their own satellites and still tie into the larger system (see Figure 2).

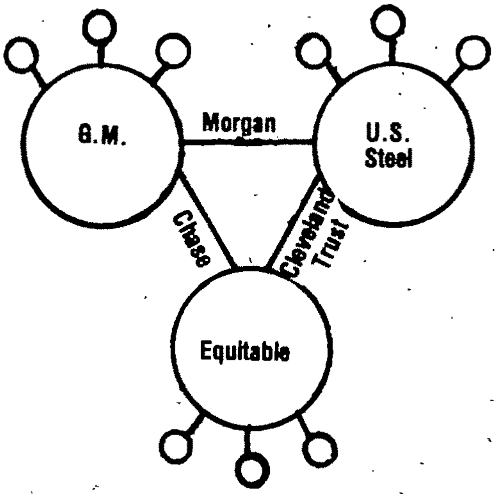


Figure 2. Reciprocity Cluster

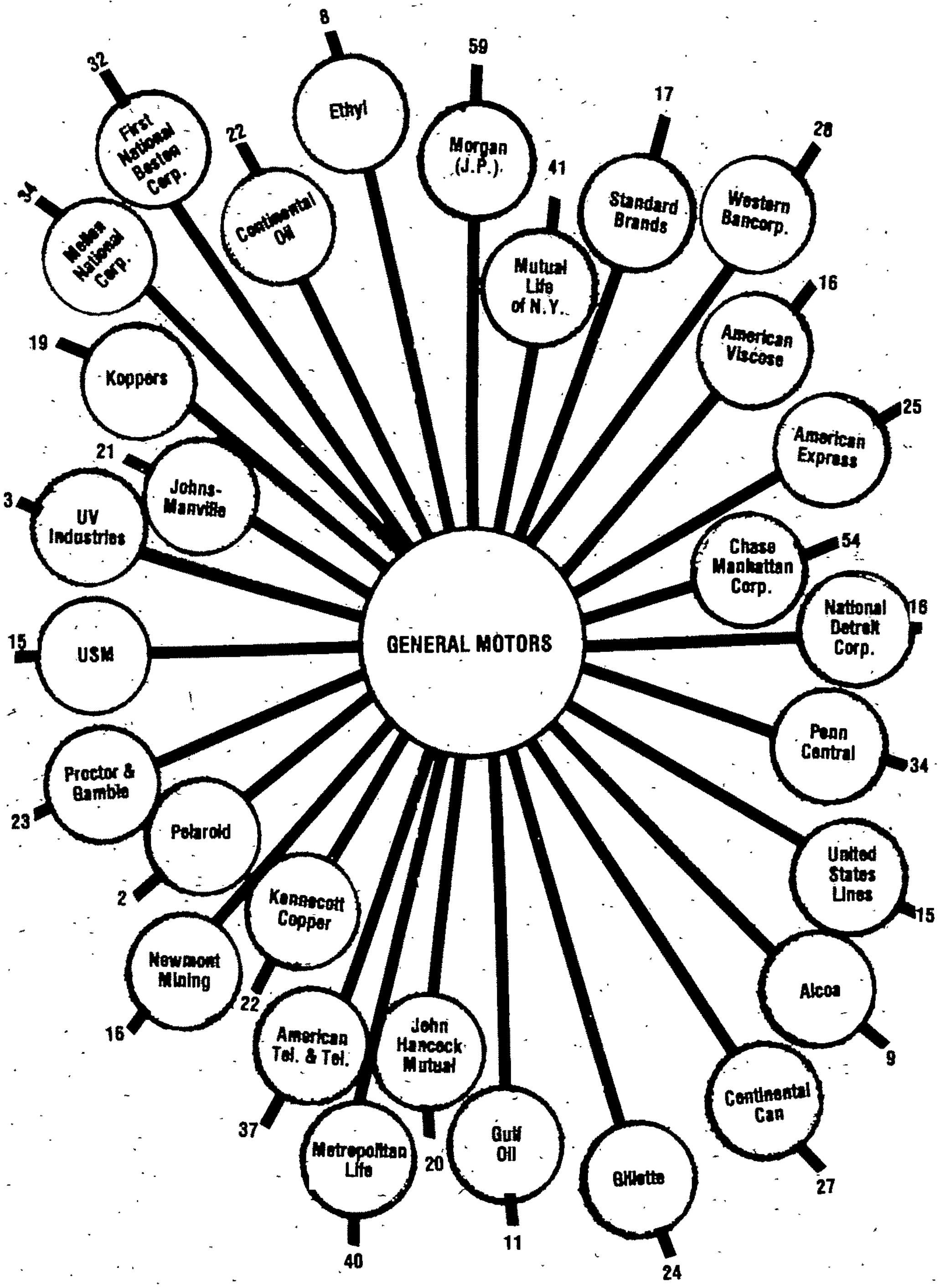
Taking these dimensions into account, coalition theory requires that the economy be organized around a series of orthogonal interests—that the auto industry, insurance industry, defense industry, and the like all find room to maneuver and organize around their interests. In order for the theory to be supported, therefore, the network should contain a variety of focal points—areas of relative density—surrounding major economic actors: Ford or General Motors (or both) at the center of an auto industry groupings; U.S. Steel organizing the heavy metals market; a major oil company (such as Exxon) at the center of still another group of firms; and a cluster of firms directly tied to a major commercial bank such as Chase Manhattan. Extensive interties among these focal points would be expected, since these might represent channels of communication and paths of potential coordination, as power blocs change and loyalties shift.

If the focal centers of the network are not multidimensional, however, this would raise serious questions about the existence of coalitional interest groups. If, for example, all insurance companies were buried in other networks, it would suggest the lack of an independent insurance interest group. Similarly, if each major steel firm was buried in an auto company network, this would indicate that steel's initiative is limited when its interest conflicted with that of the auto industry. Therefore, if the focal points of the network are dominated by a specific type of firm—all auto companies, all financial institutions, or all major utilities—this would suggest that the underlying system of coordination suppresses, rather than expresses, many potential interests.

Method

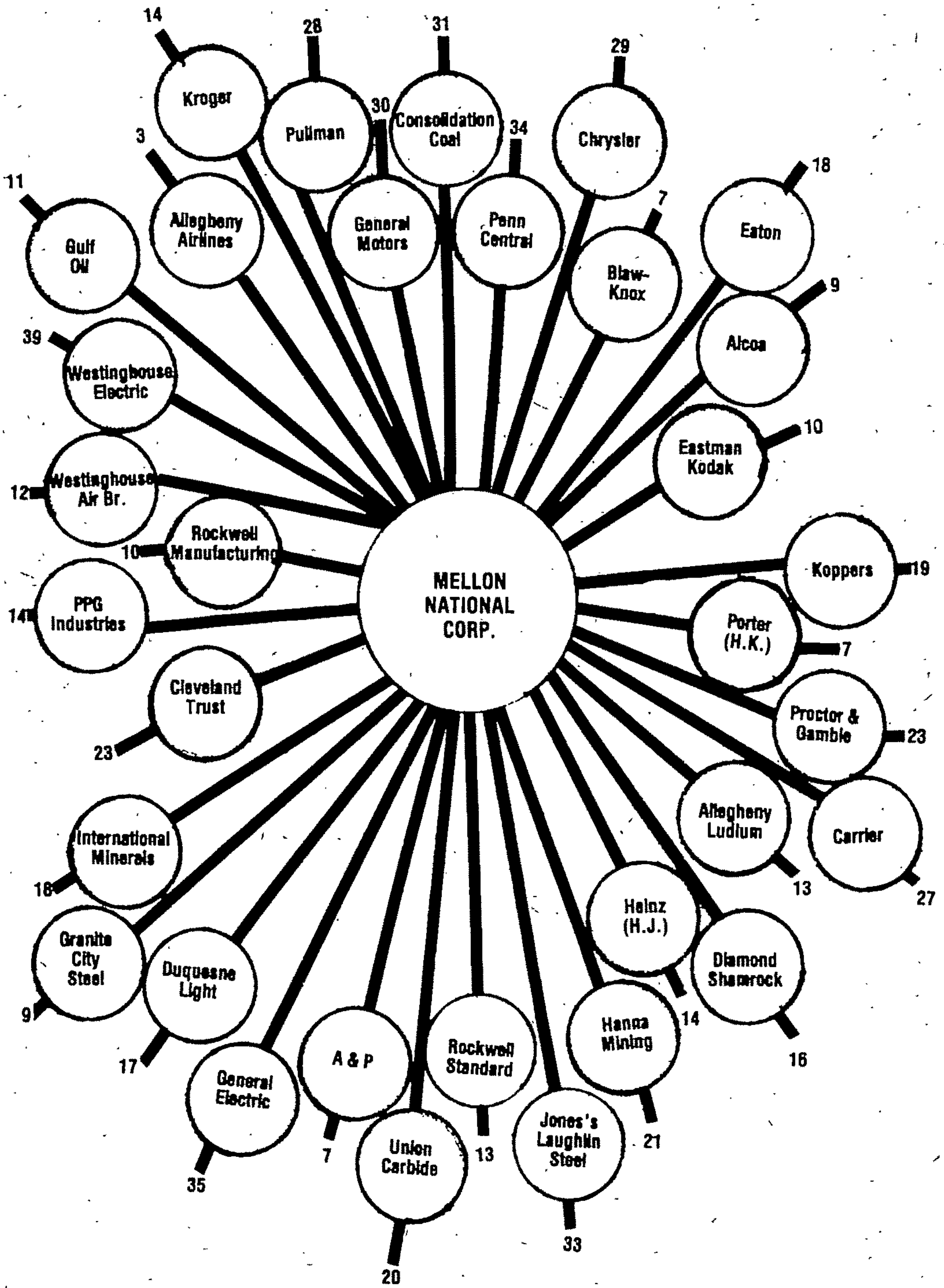
To test coalition theory, we search for the focal points of the interlock network—for those corporations which lie at the center of the densest, most interconnected regions. Such a search requires more than a simple count of the number of interlocks maintained by each firm. Consider Figures 3 and 4, which depict the network of the 1962 director interlocks for General Motors and Mellon National Bank. A simple intertie count shows that General Motors maintains direct links to 30 firms, while Mellon interlocks with 34 and, therefore, that Mellon is slightly more important. However, the corporations with which Mellon interlocks are, in turn, tied to 645 other companies, while General Motors' interties connect to 728 firms. Thus, although General Motors directly links to fewer organizations than Mellon, the network in which it is embedded is somewhat denser. Moreover, to evaluate the relative standing of the two, we must consider the strength of the connections themselves. Thus, we would judge that the many multiple interlocks maintained by Mellon indicate a greater intensity of interdependence in its network.

We must, therefore, evaluate the positional importance of each corporation on a multidimensional basis, searching for intertie patterns which reflect both the importance of the particular firm and the density of its environment. Such a measure, called "centrality," was originally developed by Bonacich (1972), modified by Mariolis (1975), and remodified by Bearden, et al. (1975). It was designed to identify pivotal corporations in the overall network—firms at the center of dense groupings but not necessarily discrete from other regions of the system. Corpo-



* The numbers next to firm names represent the total number of corporations to which the organization is intertied.

Figure 3. General Motors Interties*



* The numbers next to firm names represent the total number of corporations to which the organization is intertied.

Figure 4. Mellon Interties*

rate centrality is based on three dimensions:

(1) the number of companies with which corporation i interlocks
(2) the intensity of the interlock between corporation i and corporation j
(3) the centrality of the companies with which corporation i interlocks

$$\text{or: } C_i \sim \sum_{j=1}^{n} r_{ij} c_j \quad \text{or: } C_i = \Sigma r_{ij} c_j$$

where $r_{ii} = 0$, C_i = the centrality of corporation i, r_{ij} = the intensity of the interlock between corporation i and corporation j, c_j = the centrality of corporation j, and n = the number of corporations (Mariolis, 1975; Bearden et al., 1975).[3]

This measure is particularly useful in interlock investigations for an additional reason: intertie intensity may be manipulated to incorporate theoretical assumptions in the analysis. In this case, we define r_{ij} as $b_{ij}/d_j^{1/2}$, where b_{ij} = the number of interlocks between corporation i and corporation j, and d_j = the number of directors on the board of corporation j.[4] Thus, centrality analysis allows us to identify corporations occupying pivotal positions in the intertie network—companies in the center of the densest interlock areas (for a detailed discussion of centrality as a measure, see Mizruchi and Bunting, 1979).

Results

Table 3 presents the 20 most central corporations of the 1962 and 1966 interlock networks. A limited number of organizational types appear, with one particular sented. J. P. Morgan, the fifth largest commercial bank in the United States, was the most central organization in 1962 and the second-ranked firm of 1966. Chase Manhattan, the second largest commercial bank, ranked second in 1962 and third in 1966. The third largest insurance company, Equitable Life, was also high on both lists. In 1962, nine of the ten most central firms, and 13 of the highest 20, were financials. These organizations were not a random assortment of banks and insurance companies: they were all among the largest financials in the country, and the majority were members of the Northeastern banking establishment which has dominated corporate finance since the turn of the century.

That the areas of greatest density within the interlock network are dominated by New York financial institutions—organizations which are not likely to differ greatly in terms of overall interest—suggests a concentrated level of intercorporate power and a corresponding lack of independent action on the part of companies tied to these central units.

There is little support for coalition theory in these results: the ascendancy of financial institutions was not specific to 1962. Not only were there more banking concerns in the top ranks in 1966, they were very much the same firms, with the major New York corporate lenders remaining at the center of the densest areas in the network. This indicates a centralization and orderliness to intercorporate relationships which does not express sectoral independence. It suggests a pattern of subordination of functional clusters to an overarching integration expressed by the centrality of New York financials. Insofar as interlocks symptomize ongoing or potential joint action,

[3] Centrality scores were calculated by a series of matrix multiplications, manipulating a 1,131 × 1,131 matrix for the 1962 dataset, and a 1,111 × 1,111 matrix for 1966, solving the following equations:

$$C = RC \therefore \lambda (R - I) = 0$$

Since the equations contain one arbitrary parameter, a relative scoring scale is produced which works in the following manner: the corporation with the highest centrality is given a value of one. All other scores are normed on that assignment with the least central firm in the network given a centrality score of zero. In the 1962 interlock network, J. P. Morgan, the most central corporation in the system, was assigned a score of one. Chase Manhattan Corporation, the second most central company, scored .9003. In the present study, only ranks are reported, since this is the relevant category. See also Bearden et al. (1975), Mariolis (1975), and Mariolis, Schwartz, and Mintz (1979).

[4] Interlock intensity may also be defined as $b_{ij}/(d_i d_j)^{1/2}$ as Bearden et al. (1975) and Mariolis (1975) preferred, thus incorporating information on the size of both boards. This definition was not used in the present investigation because the number of individuals on the board of directors of the sending company was viewed as a very minor component of interlock occurrence and the additional information was not found to be of general use in the measure.

Table 3. Twenty Most Central Corporations, 1962 and 1966

Centrality Rank	Corporation	Type[a]	Fortune 500 Ranks
	1962		
1	Morgan (J P)	Bank	5
2	Chase Manhattan	Bank	2
3	Equitable Life	Ins	3
4	Chemical NY Corp	Bank	6
5	New York Life	Ins	4
6	Citicorp	Bank	3
7	Metropolitan Life	Ins	1
8	Southern Pacific	Trans	2
9	Mellon National	Bank	14
10	Manufacturers Hanover	Bank	4
11	American Tel & Teleg	Util	1
12	Penn Central	Trans	1
13	INA	Ins	†
14	Bankers Trust	Bank	9
15	General Electric	Ind	4
16	United States Steel	Ind	6
17	Westinghouse Electric	Ind	17
18	Charter New York	Bank	12
19	Harris Bancorp	Bank	23
20	Phelps Dodge	Ind	164
	1966		
1	Equitable Life	Ins	3
2	Morgan (J P)	Bank	4
3	Chase Manhattan	Bank	2
4	Chemical NY Corp	Bank	6
5	Citicorp	Bank	3
6	Metropolitan Life	Ins	2
7	American Tel & Teleg	Util	1
8	Lehman Bros	Inv	†
9	United States Steel	Ind	8
10	Westinghouse Electric	Ind	19
11	Mellon National	Bank	15
12	New York Life	Ins	4
13	United Calif Bank	Bank	14
14	General Motors	Ind	1
15	International Harvester	Ind	18
16	Pacific Mutual	Ins	33
17	Bankers Trust	Bank	7
18	First Chicago Corp	Bank	10
19	Southern Pacific	Trans	2
20	Western Bancorp	Bank	†

[a] Type abbreviations are as follows: Bank = commercial bank; Ind = industrial; Ins = life insurance company; Inv = investment bank; Mer = merchandisers, retailers, or wholesalers; Misc = miscellaneous; Trans = transportation company; Util = utility; Div = diversified investment company.

† Not ranked by *Fortune Magazine* in this year.

these findings argue against a coalition theory of market-sector groupings.[5]

It seems to us that a major objection can be made to our interpretation of financial centrality, based upon the particularities of the centrality measure as well as substantive arguments about the nature of corporate interrelationships. If we scrutinize the many types of possible network arrangements, we can identify two configurations which result in high centrality: hubs and bridges (see Figure 5). A hub is a corporation in the center of a group of interlocked firms, while a bridge is a company which links two or more hubs. Since corporate centrality is based on the centrality of all firms to which a company ties, bridges may be given high scores by our analytic technique due to their connections to a few strategically located firms. Hubs, on the other hand, obtain high centrality as a result of the cumulative weight of numerous ties to less central organizations.

It is possible, therefore, that banks and insurance companies are the most central organizations because they act as bridges connecting groups of functionally coordinated, nonfinancial corporations. In this case, the overrepresentation of financial institutions would result from measurement sensitivity to the bridging function they perform, with industrial hubs actually the focal points of actual or potential coalitions. Financial institutions would not be the organizational centers of the economy; they would be the links between these industrial groupings.

To address this possibility, we partition

[5] There is a degree of imprecision in our grouping of all financial firms into a single analytic category. In principle, there could be independent groupings representing banks, insurance companies and investment houses. Without denying the possibility, we feel that the historic relationships between New York banks and New York insurance companies (e.g., Zeitlin, 1974) suggest an enduring interdependence. Even if independence were demonstrable, it would do little to support coalition theory. The existence of two types of alliances, one organized around banking and one organized around insurance, would not allow for the full range of conflict and competition and the corresponding power limitations implied by the model.

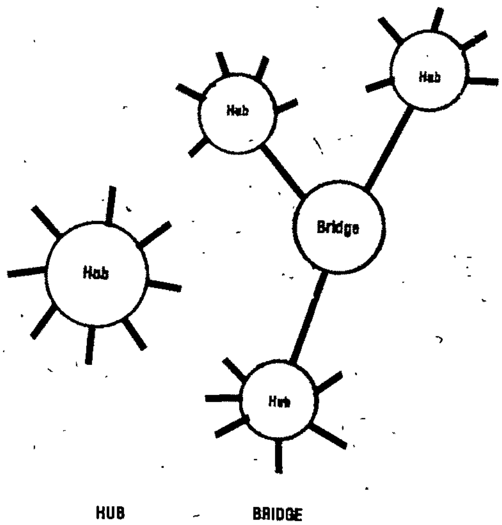


Figure 5. Network Patterns of High Centrality

the centrality of a firm into two parts: hub centrality and bridge centrality. Since a hub obtains its high ranking from interlocks with a great many firms, while a bridge obtains a large proportion of its centrality from ties to a few well-placed companies, it is possible to distinguish between the two on a systematic basis. Most centrally located corporations are linked to both large and small firms, and thus it is possible to characterize each interlock as either a "bridge" tie or a "hub" tie. This is done by calculating the proportion of a company's centrality score which is derived from its connection to a particular firm and evaluating the impact of that tie on the overall score. If one interlock contributes a disproportionate amount, or is largely responsible for the overall score of a corporation, that interlock is a "bridge tie." By subtracting the centrality gained from each bridging relationship, we can measure the degree to which a highly central corporation derives its score from its role as a hub.

A firm j is considered a significant individual influence in the environment of firm i if the centrality of j is 80% or more of the centrality of i. The contribution of such a company to the centrality score can be calculated. Thus:

$$H_i = C_i - \sum_{j=1}^{n} \propto j\,(c_j/\lambda)$$

where $\propto j = 1$ if $C_j \geq .8C_i$ and $\propto j = 0$ if $C_j < .8C_i$, H_i = hub centrality of corporation i, C_j = total centrality of corporation j, and λ = eigenvalue of solution to r matrix.
C_j = total centrality of corporation j, and λ = eigenvalue of solution to r matrix.

This procedure, then, directly addresses the possibility that financial centrality is the result of bridging functions rather than hub relationships. If financial institutions derive their centrality from their role as system bridges—linking the industrial giants of the corporate world—it will be reflected in the adjusted centrality scores produced by hub analysis. Such a result would support coalition theory.

As illustrated in Table 4, financial institutions remain the most important organizations when only hub centrality is considered. Indeed, their importance is enhanced: in 1962, two additional firms, First Chicago Bank and Mutual Life, joined the list. In both 1962 and 1966, 15 financial institutions appeared among the 20 most important hubs. In the first year, all companies which showed substantial bridging functions—AT&T, U. S. Steel, and Phelps Dodge—were nonfinancials. Of the three, Phelps Dodge was most dependent on its bridging role; in the new analysis, it dropped to rank 62.

In 1966 Continental Illinois Corporation and Harris, two Chicago area commercial banks, joined the list of most central firms. They replaced the smaller Western Bancorp and Southern Pacific, the second largest transportation company of the year. At the same time, General Motors maintained its rank, while U. S. Steel dropped from number 9 to number 20. International Harvester, on the other hand, gained position in both years. Thus, our analysis of hub centrality demonstrates that financials best maintained their positions as focal centers of the interlock system.

If we consider the degree of hubness, the importance of financials is further confirmed. Coalition theory implies that important industrial corporations would derive a large proportion of their centrality from hub relationships if they were, in fact, organizing units of the system.

Table 4. Hub Centrality, Twenty Most Central Corporations, 1962 and 1966

Centrality Rank	Corporation	Type	Fortune 500 Ranks
	1962		
1	[a] Morgan (J P)	Bank	5
2	[a] Chemical NY Corp	Bank	6
3	[a] Chase Manhattan	Bank	2
4	[a] Equitable Life	Ins	3
5	[a] Citicorp	Bank	3
6	[a] New York Life	Ins	4
7	[a] Metropolitan Life	Ins	1
8	[a] Mellon National	Bank	14
9	[a] Manufacturers Hanover	Bank	4
10	[a] Southern Pacific	Trans	2
11	[a] Bankers Trust	Bank	9
12	[a] Harris Bancorp	Bank	23
13	First Chicago Corp	Bank	10
14	[a] Penn Central	Trans	1
15	[a] Charter New York	Bank	12
16	International Harvester	Ind	21
17	[a] INA	Div	†
18	[a] Westinghouse Electric	Ind	17
19	[a] General Electric	Ind	4
20	Mutual Life	Ins	9
	1966		
1	[a] Equitable Life	Ins	3
2	[a] Morgan (J P)	Bank	4
3	[a] Citicorp	Bank	3
4	[a] Chemical NY Corp	Bank	6
5	[a] Lehman Bros	Inv	†
6	[a] Chase Manhattan	Bank	2
7	[a] Bankers Trust	Bank	7
8	[a] Mellon National	Bank	15
9	[a] International Harvester	Ind	18
10	[a] Metropolitan Life	Ins	2
11	[a] First Chicago Corp	Bank	10
12	[a] United Calif Bank	Bank	14
13	[a] Westinghouse Electric	Ind	19
14	[a] General Motors	Ind	1
15	[a] New York Life	Ins	4
16	[a] Pacific Mutual	Ins	33
17	Continental Ill Corp	Bank	8
18	[a] American Tel & Teleg	Util	1
19	Harris Bancorp	Bank	24
20	[a] United States Steel	Ind	8

[a] Included among the 20 most central corporations in the full analysis, as well.
† Not ranked by *Fortune Magazine* in this year.

This, however, is not supported by the data. The 13 financial institutions included among the original top 20 of 1962 averaged 80% hub centrality and 20% bridge centrality. That is, they derived only 20% of their centrality scores from connections to other highly central corporations. The remainder of their scores were derived from their extensive interlocking with a large number of considerably less focal companies. The nonfinancials, on the other hand, totaled 53% hub centrality; they derived almost half (47%) of their scores from connecting to other highly central corporations (mainly banks).

Thus, banks and insurance companies do not derive their high centrality from their proximity to centers of functional clusters. It seems to be the reverse: industrial firms, even highly central industrial firms, gain a substantial proportion of their centrality from their ties to bank hubs. They are themselves bridges, tying together the financial centers of the interlock network, rather than unifying satellite firms in their industrial sectors. These findings do not indicate a corporate system characterized by interorganizational alliances expressing the common interest of functionally related companies. Instead, they suggest a system organized around financial hubs which at least potentially mediate control, communication, and coordination among industries and individual firms. Coalition theory is not consistent with the shape and texture of the interlock network.

CORPORATE INTERLOCKS AND FINANCE CAPITAL THEORY

The third view of the firm, finance capital (bank control) theory, applied to the United States by Sweezy (NRC, 1939) and others, argues that the business sector is divided into autonomous interest groups, a term used to refer to "a number of corporations under common control, the locus of power being normally an investment or commercial bank or a great family fortune . . ." (Baran and Sweezy, 1966).[6]

[6] In recent years, Sweezy has abandoned the concept of interest group, arguing that developments in the post-war economy had "loosened or broken the ties that formerly bound the great interest groups together" (Baran and Sweezy, 1966:17). For a debate within the Marxist community on the role of financial institutions, see Fitch and Oppenheimer (1970), Fitch (1972), O'Connor (1968, 1972), Sweezy (1972), and Herman (1973).

Bank control theorists argue that the monopolization of lending and institutional stockholding produces the merger of financial and nonfinancial institutions into coordinated groups of corporations dominated by banking interests (Perlo, 1957; Menshikov, 1969; Fitch and Oppenheimer, 1970; Kotz, 1978).

Whereas coalition theory suggests that alliances are dissolved in the face of conflict of interest among group members, the theory of finance capital argues that conflicts are subordinated to the needs of the larger group: "the members of an interest group would naturally coordinate their policies; and in the case of conflicts, the interests of the controlling power (or of the whole group as interpreted by the controlling power) would prevail" (Baran and Sweezy, 1966).

The finance capital version of corporate behavior, therefore, argues that central financial units organize relatively autonomous, coordinated groups of member firms and maintain control over those members—that, rather than having a single national center for the system of interlocks, we should find a number of relatively discrete systems. Even with these divisions, however, banker control implies concentrated and sustained political power, since the subordination of individual interests to group needs generates unified strength. Moreover, group action organized around the needs of financial institutions suggests a community of interests among groups. To the extent, then, that financial institutions—even competing financials—share interests by virtue of their similar position vis-a-vis other types of corporations, the theory suggests that a unified group can dictate policy to other institutions and enforce its will through the coordinated political and economic action made possible by financial control.

To test this theory, we investigate the details of financial centrality in the interlock network. To support the financial interest-group model, we must find relatively discrete financial groups; and we offer two methods for measuring this. First, to evaluate the degree of integration among the most central financials, we investigate the number of direct links uniting these firms; the sparse links among them. Second, to explore whether central firms maintain exclusive ties to nonfinancial corporations, we measure the proportion of firms in the network which are linked to more than one highly central company. In this case, exclusive ties would suggest the existence of autonomous, finance-centered interest groups.

Results

Table 5 presents information on interlocking among the most central financial institutions of the 1966 network of hub centrality. The average number of ties maintained within the group is 2.4 with a range of zero (Harris Bancorp) to six (Metropolitan Life). Three companies—Mellon, Continental Illinois, and Harris Bancorp—maintained less than 10% of their potential interlocks with other system foci; four firms—Chase, Equitable, Metropolitan, and New York Life—linked to over 25% of their potential partners; the overall financial average was 17.8%. These figures demonstrate a moderate level of integration and centralization among financial institutions.

However, when we compare New York money market banks with their regional counterparts, we find a higher degree of integration among the major institutions. The five New York companies maintained an average of 22% of their potential ties, while the three Chicago banks linked to an average of 11.1%, for which First Chicago was primarily responsible. Mellon Bank (Pittsburgh) maintained 7.1%, and United California Bank tied to only 14.3% of its potential partners. The same distinction is found among insurance companies: New York firms averaged 36.4%, while Pacific Mutual, a California firm, was connected to 18.2% of possible partners.

These findings suggest that the most important financial hubs maintain direct contact with other financial institutions and that regional banks are relatively—though not absolutely—isolated from the others. This pattern argues against the existence of financially-centered interest groups, since these require isolation of the most important focal institutions from

Table 5. Interlocks Connecting Financial Focal Points, Hub Centrality, 1966

	Centrality Rank	# Ties to Other Top 20 Financials	# Possible Ties to Top 20 Financials	% of Possible Ties to Top 20[a]
Commercial banks				
NEW YORK				
Morgan (J P)	2	2	10	20.0%
Citicorp	3	2	10	20.0
Chemical	4	1	10	10.0
Chase Manhattan	6	4	10	40.0
Bankers Trust	7	2	10	20.0
		$\overline{X} = 2.2$	$\overline{X} = 10$	$\overline{X} = 22.0\%$
REGIONALS				
Mellon	8	1	14	7.1%
First Chicago	11	3	12	25.0
United California	12	2	14	14.3
Continental Ill	17	1	12	8.3
Harris	19	0	12	0
		$\overline{X} = 1.4$	$\overline{X} = 12.8$	$\overline{X} = 10.9\%$
Insurance companies				
NEW YORK				
Equitable	1	3	11	27.3%
Metropolitan	10	6	11	54.5
New York Life	15	3	11	27.3
		$\overline{X} = 4$	$\overline{X} = 11$	$\overline{X} = 36.4\%$
REGIONAL				
Pacific Mutual	16	2	11	18.2%
Investment banks				
Lehman Bros	5	5	14	35.7%
		Total $\overline{X} = 2.4$	Total $\overline{X} = 11$	Total $\overline{X} = 17.8\%$

[a] Interlocks among competitors are prohibited by federal law. For commercial banks, this legal prohibition applies only to other banks operating in the same locality.

each other; and it differs from previous interlock analyses, which have viewed New York as one of several densely interlocked subsets within the larger system (see Dooley, 1969; Allen, 1978; Sonquist and Koenig, 1975). We interpret these results as suggesting that New York is the base of a national network of corporate interlocks, uniting regional clusters into a loosely integrated whole.[7]

This evidence could still fit a modified version of bank control theory if each of the central financial institutions maintained exclusive ties to many noncentral industrials in its domain. In this case, the dense interlocks among focal units would indicate temporary alliances, which, when broken, would return the system to competing financial empires. A general impression of the extent to which distinct ties are maintained is provided by noting that the most central financials linked to 678 companies in all, of which 322 (47.6%) were nonunique—almost one-half of the firms involved tied to more than one of the focal centers.

Table 6 provides a more detailed breakdown of shared interlocks. A pattern of financial interest groups would be suggested by a great many zero cells in the table, with other entries indicating large numbers of shared ties. This configuration

[7] This interpretation is supported by the structure of interregional linkages. The 20 most central companies of the 1966 network are located in five areas: New York, Chicago, Pittsburgh, Detroit, and Los Angeles. Only two regions—New York and Pittsburgh—tie to all locations. Chicago, Los Angeles, and Detroit do not connect to each other. Pittsburgh, we believe, is not a national hub. Its ties to other regions derive mainly from Westinghouse Electric's links to nine other central companies in four locations. Mellon, the main Pittsburgh commercial bank, interlocks with a total of four relevant companies and connects to New York and Chicago only. Its integrative role is mainly on the regional level, and it is tied into the national system principally through its connections to Westinghouse Electric. Thus, with some exceptions, the regions relate "upward" to New York and do not tie to each other directly. (For a fuller discussion of this pattern, see Mintz and Schwartz, forthcoming).

Table 6. Matrix of Shared Directorship Ties Among the Most Central Financials, Hub Centrality, 1966

	Morgan	Citicorp	Chem'l	Lehman	Chase	Bankers Trust	Mellon	Metro Life	First Chi'go	Unt'd Calif.	N.Y. Life	Pac. Life	Contin'l Ill.	Harris
Equitable	12	8	18	6	17	5	8	7	2	2	11	4	7	3
Morgan		14	10	13	0	8	7	10	3	3	7	2	5	1
Citicorp			6	7	14	4	5	12	4	4	10	5	1	1
Chemical				5	6	5	1	6	5	1	10	1	4	3
Lehman					5	7	7	11	6	8	7	12	1	1
Chase						8	6	9	4	2	7	2	1	5
Bankers Trust							2	7	3	2	4	3	2	1
Mellon								6	2	0	2	3	2	2
Metropolitan Life									3	3	10	4	2	1
First Chicago										2	7	1	8	8
United California											3	16	2	1
New York Life												3	3	3
Pacific Mutual													1	1
Continental Illinois														6

does not occur. Instead, there is a very strong pattern of overlapping ties, with the average pair of top financials linking to five of the same less central corporations. Moreover, among the 105 pairings in the table, only two do not share at least one interlock. A vast majority of the most central institutions maintained mutual interlocks with each of the other focal firms.

These results indicate that by-and-large, the less central corporations in the system do not cluster around specific focal organizations. Instead, the overwhelming tendency is for the less central firms to be tied to several different hubs. There is, of course, much variation among noncentral corporations. Consolidated Edison, for example, interlocked with seven of the financial hubs; Allied Chemical and Chrysler tied to five; Kellogg to one. In general, the larger, more important companies linked to a greater number of central organizations. Peripheral, smaller firms tied to fewer of the central units. Here again, the pattern is counter to the existence of separate clusters, since meaningful groups would allocate important industrials into the exclusive domain of central dominant financial firms.

These findings appear to contrast sharply with previous interlock studies which have identified corporate groupings. Dooley (1969), Sonquist and Koenig (1975), and Allen (1978) have used intertie patterns to identify and analyze financially-centered regional clusters (See also Knowles, 1973; Menshikov, 1969; Pelton, 1970). We argue, however, that previous research, when carefully scrutinized, confirms the results reported here, since groups identified in earlier studies were neither discrete nor coordinated; they were actually unsegregated areas of relative density in an otherwise integrated network. Moreover, earlier research has found financial institutions to be the key units of the interlock network (Levine, 1972; Allen, 1974; Bearden et al., 1975; Mintz and Schwartz, 1981) and although the evidence is often left unanalyzed, most of these studies find a unique set of relationships in the New York area. This reinforces our contentions that (1) New York banks and insurance companies are the focal institutions of an integrated national network, and that (2) other sectors of the network are mediated by local financial institutions.

CONCLUSION

The evidence presented here suggests that the large American corporation is not an autonomous unit, that interorganizational alliances are not characterized by industry coalitions, and that discrete financial groups are not the typical unit of corporate organization. We have found that the

interlock network is dominated by a handful of major New York banks and insurance companies. Since densely interlocked companies carry the potential to coordinate intercorporate action, these institutions are prime candidates for that role. That they maintain their position in various analyses underscores this point.

These findings are inconsistent with the three models of corporate interaction explored above. They are consistent, however, with a modified finance capital theory which locates the importance of banks and insurance companies in their domination of capital flows and not in discrete spheres of influence. The shape of the network suggests a corporate structure loosely organized around the needs of financial institutions—major New York money market firms, in particular—which are themselves intertied and increasingly interrelated. This points to a centralized structure of intercorporate relationships and suggests a process which minimizes corporate conflict. These traces of unity indicate that corporate power is not dispersed, that large companies are indeed involved in each other's affairs, and that the product of this involvement is an inhibiting effect on the mobilization of individual interests. The overall shape of the network is consistent with the findings of Domhoff (1970, 1978), Ratcliff (1979–1980, 1980), Useem (1978, 1979), Whitt (1982), and others that the corporate community is capable of coordinated economic decision making and united political action.

REFERENCES

Aldrich, Howard
1979 Organization and Environments. New Jersey: Prentice Hall.

Aldrich, Howard and Jeffrey Pfeffer
1976 "Environments of organizations." Pp. 79–105 in A. Inkeles (ed.), Annual Review of Sociology, Vol. 11. Palo Alto: Annual Reviews.

Allen, Michael P.
1974 "The structure of interorganizational elite cooptation: interlocking corporate directors." American Sociological Review 39:393–406.
1978 "Economic interest groups and the corporate elite structure." Social Science Quarterly 58:597–615.

Baran, Paul and Paul Sweezy
1966 Monopoly Capital. New York: Monthly Review Press.

Bearden, J., W. Atwood, P. Freitag, C. Hendricks, B. Mintz, and M. Schwartz
1975 "The nature and extent of bank centrality in corporate networks." Paper presented at the Annual Meeting of the American Sociological Association, San Francisco.

Bell, Daniel
1961 The End of Ideology. New York: Collier.

Berkowitz, Stephen D., Paul J. Carrington, Yehuda Kotowitz, and Leonard Wavarman.
1979 "The determination of enterprise groupings through combined ownership and directorship ties." Social Networks 1:415–35.

Berle, A. and G. Means
1932 The Modern Corporation and Private Property. New York: Harcourt, Brace and World.

Bonacich, P.
1972 "Technique for analyzing overlapping memberships." Sociological Methodology. San Francisco: Jossey-Bass.

Brandeis, Louis
1914 Other People's Money. New York: Frederick A. Stokes Co.

Bunting, David
1976–1977 "Corporate interlocking." The Journal of Corporate Action 1:6–15; 2:27–37; 3:4–11; 4:39–47.

Bunting, David and Jeffrey Barbour
1971 "Interlocking directorates in large American corporations, 1896–1962." Business History Review 45:317–35.

Burt, Ronald
1979 "A structural theory of interlocking corporate directorates." Social Networks 1:415–35.
1980 "On the functional form of corporate cooptation: empirical findings linking the intensity of market constraint with the frequency of directorate ties." Social Science Research 9:146–77.

Dahl, Robert
1961 Who Governs? Democracy and Power in an American City. New Haven: Yale University Press.

Directory of Directors of New York
Annually New York: Directory of Directors Company, Inc.

Domhoff, G. William
1970 The Higher Circles. New York: Random House.
1978 The Powers That Be. New York: Random House.

Dooley, Peter
1969 "The interlocking directorate." American Economic Review 59:314–23.

Dun and Bradstreet
Annually Dun and Bradstreet Million Dollar Directory. New York: Dun and Bradstreet Inc.

Federal Trade Commission
1951 Report on Interlocking Directorates. Washington, D. C.: U.S. Government Printing Office.

Fennema, Meindert
Forthcoming International Networks of Banks and Industry: 1970–1976.
Fennema, Meindert and B. Schijf
1979 "Analyzing interlocking directorates: theory and method." Social Networks 1:297–332.
Fitch, Robert
1972 "Sweezy and corporate fetishism." Socialist Revolution 12:93–127.
Fitch, Robert and Mary Oppenheimer
1970 "Who rules the corporations?" Socialist Revolution 4:73–108; 5:61–114; 6:33–94.
Fortune Magazine
Bi-weekly Chicago: Time, Inc.
Galbraith, John K.
1967 The New Industrial State. New York: New American Library.
Gogel, Robert
1976 "Interlocking directorships and the American corporate network." Ph.D. Thesis. Department of Sociology, University of California, Santa Barbara.
Gogel, Robert and Thomas Koenig
1981 "Commercial banks, interlocking directorates and economic power: an analysis of the primary metals industry." Social Problems. In press.
Herman, Edward
1973 "Do bankers control corporations?" Monthly Review 25:12–29.
Hunter, Floyd
1953 Community Power Structure. Chapel Hill: The University of North Carolina Press.
Kaysen, Carl
1957 "The social significance of the modern corporation." American Economic Review 47:311–9.
Knowles, James
1973 "The Rockefeller financial group." Warner Modular Publications. Module 343:1–59.
Koenig, Thomas and Robert Gogel
1981 "Interlocking directorates as a social network." American Journal of Economics and Sociology 40:37–50.
Kotz, David
1978 Bank Control of Large Corporations in the United States. Berkeley: University of California Press.
Larner, Robert J.
1970 Management Control and the Large Corporation. New York: Dunellen Publishing.
Levine, Joel
1972 "The sphere of influence." American Sociological Review 37:14–27.
1977 "The network of corporate interlocks in the United States." Paper presented at the meetings of the American Sociological Association, Chicago.
Lundberg, Ferdinand
1968 The Rich and the Super Rich. New York: Lyle Stuart.
Mace, Myles
1971 Directors: Myth and Reality. Cambridge: Harvard University Press.
Mariolis, Peter
1975 "Interlocking directorates and the control of corporations." Social Science Quarterly 56:425–39.
1977 "Type of corporation, size of firms and interlocking directorates: a reply to Levin." Social Science Quarterly 58:511–3.
1978 "Bank and financial control among large U. S. corporations." Department of Sociology, Ph.D. dissertation. State University of New York at Stony Brook.
Mariolis, Peter, Michael Schwartz, and Beth Mintz
1979 "Centrality analysis: a methodology for social networks." Paper delivered to meeting of American Sociological Association, Boston.
Menshikov, Sergei
1969 Millionaires and Managers. Moscow: Progress Publishers.
Mills, C. Wright
1956 The Power Elite. New York: Oxford University Press.
Mintz, Beth
1978 "Who controls the corporation: a study of interlocking directorates." Department of Sociology, Ph.D. Dissertation. State University of New York at Stony Brook.
Mintz, Beth and Michael Schwartz
1981 "The structure of power in American business." Social Problems. In press.
Forthcoming The Structure of Power in the American Intercorporate System.
Mizruchi, Mark
1980 "The structure of the American corporate network: 1904–1974." Department of Sociology, Ph.D. Dissertation. State University of New York at Stony Brook.
1982 The Structure of the American Corporate Network: 1904–1974. Beverly Hills: Sage.
Mizruchi, Mark and David Bunting
1979 "Influence in corporate networks: an examination of four measures." Paper presented at the meetings of the American Sociological Association, Boston.
Mokken, R. J. and F. N. Stokman
1974a "Traces of power I and II: power influence as political phenomenon." Paper presented at the conference of the European Consortium for Political Research, Strasbourg.
1974b "Interlocking directorates between large corporations, banks and other financial companies and institutions in the Netherlands in 1969." Paper presented at the conference of the European Consortium for Political Research, Strasbourg.
1979 "Corporate-government networks in the Netherlands." Social Networks 1:333–58.
Moody's Bank and Financial Manual, Industrial Manual, Public Utility Manual, Transportation Manual
Annually New York: Moody's Investor Service, Inc.
Moore, Gwen
1979 "The structure of a national elite network." American Sociological Review 44:673–91.

National Resources Committee
1939 "The structure of the American economy." Washington, D. C.: Government Printing Office. Reprinted in Paul M. Sweezy (1953), The Present as History. New York: Monthly Review Press.

Norich, Sam
1980 "Interlocking directorates, the control of large corporations, and patterns of accumulation in the capitalist class." Pp. 83–106 in M. Zeitlin (ed.), Classes, Class Conflict and the State. Cambridge, Mass.: Winthrop.

O'Connor, James
1968 "Finance capital or corporate capital?" Monthly Review 20:30–5.
1972 "Question: who rules the corporation? Answer: the ruling class." Socialist Revolution 7:117–50.

Palmer, Donald
1980a "Broken ties: a theoretical and methodological critique of corporate interlock research." Ph.D. Dissertation. Department of Sociology, State University of New York at Stony Brook.
1980b "Broken ties: some political and interorganizational determinants of interlocking directorates among large American corporations." Paper presented at the meetings of the American Sociological Association, New York.

Pelton, Richard
1970 Who Really Rules America? Somerville: New England Free Press.

Pennings, Johannes
1980 Interlocking Directorates. San Francisco: Jossey-Bass.

Perlo, Victor
1957 The Empire of High Finance. New York: International Publishers.

Pfeffer, Jeffrey and Gerald Salancik
1978 The External Control of Organizations: A Resource Dependence Perspective. New York: Harper and Row.

Pujo Committee
1913 Investigation of Concentration of Control of Money and Credit. Washington, D. C.: Government Printing Office.

Ratcliff, Richard
1979–1980 "Capitalist class structure and the decline of older industrial cities." Insurgent Sociologist 2–3:60–74.
1980 "Banks and the command of capital flows: an analysis of capitalist class structure and mortgage disinvestment in a metropolitan area." Pp. 107–32 in M. Zeitlin (ed.), Classes, Class Conflict and the State. Cambridge, Mass.: Winthrop Publishers.

Ratcliff, Richard, Mary Elizabeth Gallagher, and Kathryn Strother Ratcliff.
1979 "The civic involvement of bankers: an analysis of the influence of economic power and social prominence in the command of civic policy positions." Social Problems 26:298–313.

Rochester, Anna
1936 Rulers of America. New York: International Publishers.

Rose, Arnold
1967 The Power Structure. New York: Oxford University Press.

Scott, John
1979 Corporations, Classes and Capitalism. London: Hutchinson University Library.

Scott, John and Michael Hughes
1976 "Ownership and control in a satellite economy: a discussion from Scottish data." Sociology 10:21–41.

Seybold, Peter
1978 "The development of American political sociology." Ph.D. Dissertation. Department of Sociology, State University of New York at Stony Brook.

Sonquist, John and Thomas Koenig
1975 "Interlocking directorates in the top United States corporations: a graph theory approach." The Insurgent Sociologist 5:196–230.

Soref, Michael
1979 "Research on interlocking directorates: an introduction and a bibliography of North American sources." Connections 2:84–6.

Standard and Poor's Register of Corporations, Directors and Executives
Annually New York: Standard and Poor's Corporation.

Sweezy, Paul
1972 "The resurgence of finance capital: fact or fancy." Socialist Revolution 8:157–90.

Truman, David
1951 The Government Process. New York: Random House.

United States Congress, House of Representatives
1965 "Interlocks in corporate management." Washington, D. C.: Government Printing Office.
1968 "Commercial banks and their trust activities: emerging influence on the American economy." Washington, D. C.: Government Printing Office.

Useem, Michael
1978 "The inner group of the American capitalist class." Social Problems 25:225–40.
1979 "The social organization of the American business class." American Sociological Review 44:553–71.

Warner, W. Lloyd and Darab Unwalla
1967 "The system of interlocking directorates." Pp. 121–37 in Warner, Unwalla, and Trimm (eds.) The Emergent American Society: Vol. 1: Large Scale Organizations. New Haven: Yale University Press.

Whitt, J. Allen
1976 "Means of movements: the politics of modern transportation systems." Ph.D. Dissertation. Department of Sociology, University of California, Santa Barbara.
1982 Means of Motion. Princeton: Princeton University Press.

Williamson, O. E.
1965 "A dynamic theory of interfirm behavior." Pp. 27–43 in Evan (ed.), Interorganizational

Relations. Philadelphia: University of Pennsylvania Press.
Yago, Glenn
1980 "Corporate power and urban transportation: a comparison of public transit's decline in the United States and Germany." Pp. 296–323 in M. Zeitlin (ed.) Classes, Class Conflict and the State. Cambridge, Mass.: Winthrop Publishers.
Zeitlin, Maurice
1974 "Corporate ownership and control: the large corporation and the capitalist class." American Journal of Sociology 79:1073–119.

TAXONOMIES OF INSTITUTIONAL STRUCTURE: DUAL ECONOMY RECONSIDERED*

LYNNE G. ZUCKER
CAROLYN ROSENSTEIN
University of California, Los Angeles

Hypotheses drawn from the dual economy approach concerning the effects of institutional differentiation of the economy on labor market conditions and labor force characteristics are examined for each of four major taxonomies using data from the NORC General Social Surveys. Mean labor market conditions and labor force characteristics show inconconsistent patterns of difference between sectors across the different taxonomies and do not support hypothesized sectoral differences. Income attainment and poverty earnings within sectors are consistently related to labor market conditions, but labor force characteristics show a much less consistent pattern across taxonomies. Only one of the four taxonomies produces overall significant differences between sectors (for predicting earnings only). Generally, the four taxonomies produce inconsistent patterns of results and show a mixed pattern of support for dual economy predictions.

Investigations of labor market processes have emphasized the impact of individual attributes on attainment outcomes (e.g., Sewell and Hauser, 1975; Duncan et al., 1972; Blau and Duncan, 1967). In reaction to these individualistic models of status attainment, sociologists have recently adopted more structural approaches (for a review, see Zucker, 1979), in many cases borrowing directly from the economics literature. Structural arguments in economics have taken two forms: dual economy theory, which focuses on differences between firms and/or industries (Averitt, 1968; O'Connor, 1973); and dual labor market theory, which focuses on the consequences of labor market segmentation while largely ignoring industrial differentiation (Piore, 1973, 1975). The emerging structural approach in sociology rests primarily on the dual economy perspective (e.g., Bibb and Form, 1977; Baron and Bielby, 1980), with some aspects of the dual labor market approach loosely borrowed (e.g., Beck et al., 1978; Tolbert et al., 1980).

* Direct all correspondence to Lynne G. Zucker, Department of Sociology, University of California, Los Angeles, CA 90024.

The first author is responsible for the basic framework, development of the approach, and drafting of the manuscript; the second author is responsible for reconstructing the Beck et al. (1978) data set and conducting the analysis. We appreciate the information provided by E. M. Beck, Patrick M. Horan, and Charles M. Tolbert II, which enabled us to reconstruct their data set. We are grateful to Valerie Oppenheimer and Melvin Seeman for their extensive comments on an earlier draft, to Phillip Bonacich for advice on the data analysis, and to the anonymous reviewers for their helpful comments on revision. The research was supported by a grant from the Research Committee of the Academic Senate, UCLA, to Lynne G. Zucker. We wish to thank the Data Archives Library, Institute for Social Science Research, UCLA, and the Inter-university Consortium for Political and Social Research for making the data available. The data were originally collected by the National Opinion Research Center.

There is no unitary approach explicating the causes and consequences of the dual structure of the economy (Oster, 1979:33). However, broad comparisons among the different approaches are possible because of their common focus on economic differentiation at the institutional level, separating one type of firm or industry from another (Averitt, 1968; Bluestone et al., 1973). When the causes of this differentiation are investigated, the level of analysis is typically the firm (e.g., Averitt, 1968): some large firms, constituting the core sector, dominate the market (monopolistic or oligopolistic firms), while other smaller firms, constituting the periphery, have much less economic control (competitive capitalism). However, there is little consensus on the underlying causes of this differentiation. For example, elements such as industry profit or type of support (private versus governmental) are central to some definitions, peripheral or omitted in others. Further, when multiple causes are identified—such as industry concentration, size, and profit—they are generally not highly correlated, or show inconsistent patterns of relationship, with each other and with labor market indicators (see Oster, 1979; Tolbert et al., 1980; Wallace and Kalleberg, 1981; Kaufman et al., 1981).

However, most empirical work in the dual economy tradition focuses on the consequences of differentiation—the effects of the industrial structure on labor market conditions and on the individual worker (Bluestone et al., 1973; Bibb and Form, 1977; Beck et al., 1978; Tolbert et al., 1980). In this case, it is generally expected that more global institutional differences will suffice, and the level of analysis is typically at the industry level (e.g., Beck et al., 1978).[1] Differences between industrial sectors are expected in labor market conditions, such as unemployment and unionization (Bluestone et al., 1973), and in labor force characteristics, such as sex and education (Bibb and Form, 1977). Yet there are few links between the theoretical writing on dual economy and differentiation of labor market and labor force characteristics. Those which do exist explicitly use concepts borrowed from the dual labor market approach (Hodson, 1978:433–4; and Edwards, 1979:ch. 9), though the correspondence between primary labor market (with high-wage jobs, good working conditions, and internal labor markets) and core sector, or between secondary labor market and periphery, is far from clear (only Edwards, 1975:21, argues explicitly for an isomorphic relationship).

Industrial differentiation is also expected to have a substantial impact on workers' income attainment and on the probability that workers have poverty-level earnings. Translated into working hypotheses in the dual economy literature, special emphasis is placed on "low-wage industries" and on the differential placement of workers within them (see especially Bluestone et al., 1973:45–7, 101–3, 164–6). However, explanations of the persistence of the "working poor" in several of the most important empirical studies of dual economy (Bluestone et al., 1973:ch. 2; Beck et al., 1978) have been derived in part from dual labor-market models (Gordon, 1972:ch. 6; Doeringer and Piore, 1975; Vietorisz and Harrison, 1973). In addition, the few explicit hypotheses offered concerning overall income attainment also incorporate the dual labor-market approach and findings. Results from research on dual labor markets have been used to generate these hypotheses. For example, human capital variables have been found to be more important in determining income in the primary, as compared to the secondary, labor market (Osterman, 1975, as cited in Beck et al., 1978:705; Doeringer and Piore, 1975:72; Rosenberg, 1976:1, 12–3, 21); hypotheses concerning core and periphery explicitly rest on these findings (Beck et al., 1978).

In general, it is expected that income attained in the secondary labor market will not depend as heavily on individual attributes (Gordon, 1972:50–1). Some supportive evidence has been reported (Bibb and Form, 1977) demonstrating, for

[1] In part this decision may rest on the availability of data at the industry, not the firm, level. Some have explicitly argued for analysis of firms even when consequences of differentiation are investigated (e.g., Baron and Bielby, 1980).

example, that income attained by blue-collar workers depends more on structural variables (e.g., firm size, occupational prestige) than on individual attainment (e.g., education). Other studies have produced conflicting evidence concerning the effect of education on income in the different sectors or markets (compare results in Osterman, 1975, to those in Bluestone et al., 1973:88–90).

The varied emphasis on different causes and consequences—and their linkage—produces at least a slightly different taxonomic division of industries into "dual" sectors in each empirical study. At least a dozen different sector taxonomies have been proposed.[2] Most divide industries into two sectors, for simplicity labeled core/periphery here (elsewhere termed center/periphery or monopoly/competitive). In one case, a state sector is added (Hodson, 1977).

Four operational definitions of sector (see Appendix), all serving as benchmarks in recent dual economy research (see, for example, Wallace and Kalleberg, 1981; Kalleberg et al., 1980; Kaufman et al., 1981; Pfeffer and Ross, 1979), are investigated in more detail here: two developed by the Beck, Horan, and Tolbert group (1978 and 1980), one by Bibb and Form (1977), and one by Hodson (1977).[3] Each of these four taxonomies assigns industries to sectors differently (refer again to Appendix), contrary to Beck et al.'s (1978:710) contention that there is "consensus on the sectoral location of the majority of industries." Comparing the 53 industrial categories of the Beck-Horan-Tolbert taxonomy with those used in the other taxonomies produces different levels of agreement: 83% with Bibb-Form and 57% with Hodson. The Bibb-Form and Hodson taxonomies agree on only 58% of the assignments. Comparing the more complex (86-category) Tolbert-Horan-Beck (1980) classification, there is consensus on about two-thirds of the industries with each of the other three taxonomies.

Most recent re-examinations of dual economy taxonomies have not attempted to replicate earlier work (e.g., Wallace and Kalleberg, 1981); that is, the variables and analytical procedures selected have been dissimilar. In contrast, the variables and procedures used in the present paper are selected because of their importance in earlier work. However, because each of these four definitions of sector involve somewhat different variables and procedures, those used by the Beck, Horan, and Tolbert (1978) group are followed most closely here for uniformity.

Using a sample drawn from the National Opinion Research Center's General Social Survey, two empirical questions are addressed: (1) Do differences (sometimes very small) in classification of industries across dual economy taxonomies lead to significant differences in empirical analysis? (2) What is the extent of support for the major theoretical predictions concerning the effects of industrial sector on labor market conditions and labor force characteristics? Using the natural log of income and proportion with poverty earnings as the dependent variables, the results obtained from the four different classifications are compared in terms of their labor market and labor force characteristics and in terms of processes of earnings determination. General conclusions are drawn about the sensitivity of empirical analysis to differences in definition of economic sectors. The usefulness of each taxonomy for exploring the effects of the institutional structure on individual economic outcomes is discussed.

SAMPLE, VARIABLES, AND PROCEDURES

To compare the four taxonomies of industrial sectors empirically, data are ob-

[2] Not only are the explanatory variables different, but also the actual procedure followed in defining sectors is quite varied: (1) industries are divided according to rough judgments; typical criteria are stability of employment or wage rates (e.g., Bibb and Form, 1977); (2) empirical criteria related to institutional structure are developed, but criteria for actual assignment of industries to sectors are not specified (e.g., Bluestone et al., 1973; Beck et al., 1978; Hodson, 1977); and (3) sectors are defined empirically, using different sets of industrial, labor market, and/or labor force variables (Tolbert et al., 1980; Oster, 1979; Pfeffer and Ross, 1979; Kaufman et al., 1981).

[3] Tolbert et al. (1980) also develop a continuous model which is not investigated here. It did not produce results as strong as those obtained with the dichotomous model (see especially Table 8, p. 1114).

tained from the National Opinion Research Center's General Social Surveys for 1975 and 1976 (Davis, 1975, 1976), with a total sample of 1,584, composed of those working full-time, part-time, or looking for work.[4] The Beck-Horan-Tolbert procedures are followed, using the same independent and dependent variables (see Beck et al., 1978, for a complete description), with three important changes: (1) Ninety-nine nonearners were deleted from the sample in all four taxonomies, in response to Hauser's (1980) critique of the Beck, Horan, and Tolbert procedures, because the latter had treated some job changers and persons out of the labor force as if they had worked but had no income;[5] (2) parental occupational prestige was deleted from the equations because it had "trivial negative effects" (Hauser, 1980:707); and (3) logistic regression was used for analyses using the dichotomous dependent variable, proportion with poverty earnings. The methods of analysis are discussed in more detail below.

The Beck et al. (1978:708–9) definition of variables is used. See footnotes c to n, Table 1, in the present paper for replicated and modified definitions of variables. Specific hypotheses are introduced below in the discussion of findings. Assembling hypotheses to guide the analysis of the effects of dual economy on labor market and labor force characteristics is a frustrating task. What has been said of the dual labor market approach applies with equal force to the recent empirical work on the dual economy (Cain, 1976:221): "[The] theories are sketchy, vague, and diverse if not internally conflicting. Description, narratives, and taxonomies crowd out model development." Further indicating the lack of agreement in this area, variables thought by some researchers to be unrelated to dual economy differentiation, such as occupational prestige (Oster, 1979), are used by others (see its use in Beck et al., 1978, borrowing from Osterman, 1975, and Freiman, 1976).

In the next section, specific hypotheses are investigated in the context of analyzing three important consequences of sector differentiation: labor market conditions and labor force characteristics in core and periphery; relative importance of a given variable cluster (labor market conditions or labor force characteristics) in predicting annual income and proportion of workers with poverty level incomes; and extent to which effects of the variables differ between the sectors of each taxonomy.

COMPARISON OF SECTOR CHARACTERISTICS

The mean labor market conditions and labor force characteristics in core and periphery (and, in Hodson, state) are presented in Table 1. When the means for each taxonomy are examined, it is clear that the degree and direction of significant differences between sectors are inconsistent.[6]

Eliminating Hodson because of the overall lack of significant differences, only

---

[4] The same sample is used to permit direct comparison with the Beck, Horan, and Tolbert work, though larger data sets are now available. We also used the full sample available from the General Social Survey (GSS), which contained all but one of the variables used in the analysis (1974 through 1978) to be sure that the results obtained were not an artifact of the small sample size. The results were remarkably consistent. No differences between the samples were significant; the same direction and magnitude of difference between sectors was uniformly found.

[5] The error in sample selection was (Hauser, 1980:711) "compounded by using a logarithmic transformation of earnings, which gives a great deal of importance to observations at or near the zero point of the transformed variable. Because of this error, [Beck et al.'s] analysis of earnings (but not of labor force composition) are wrong in important respects." Findings reported in Tables 1 and 2 for our replication of the Beck-Horan-Tolbert taxonomy, therefore, should be compared to Hauser's (1980, Table 1 and Table 3), as well as to Beck et al.'s (1978). Slight discrepancies result from some methodological differences, such as Hauser's use of pairwise deletion for several variables contrasted with the assignment of values to missing cases done in the analysis reported here. For other differences in specific assignments, compare our assignments (see footnotes to Table 1) with Hauser's (1980:705 and footnotes to Table 1).

[6] The hypotheses tested are often not directional, or not uniformly predicted to be in the same direction. It has been customary, therefore, to use a .10 level of significance for the hypotheses in most investigations of dual economy theory (Beck et al., 1978; Hauser, 1980).

half of the variables in Table 1 have consistently significant differences between core and periphery means. Further, the inconsistencies often occur for variables highlighted as critical differentiators of core and periphery (e.g., work stability and educational credentialing). There are also reversals of the direction of significant differences across classifications for nearly half of the variables, so that a higher proportion is found in the core sector in one taxonomy, but in the periphery in another. For example, years of schooling is significantly lower in the periphery in Beck-Horan-Tolbert, but significantly higher in the periphery for Bibb-Form.

Posing an even more significant challenge to these operationalizations of dual economy, the differences in means between sectors within each taxonomy are frequently not consistent with the theoretical predictions.

Dependent Variables. Workers in the core sector are expected to earn more income; and correspondingly, the core sector should have a lower proportion of workers receiving income below the poverty level. The findings support these expectations. For the Tolbert-Horan-Beck taxonomy, this result may in part reflect the circularity that arises from their use of income (and other characteristics of the work force) in deriving sectors (see Hodson and Kaufman, 1981; Horan et al., 1981).

Labor Market Conditions. A higher percentage of union membership is predicted in the core because unions are larger and more common in this sector than in the periphery, in part because of the larger size of core firms (Averitt, 1968:ch.7) and the need for stability in the work force (derived from Reich et al., 1973).[7] Stability of employment, both in terms of current and past unemployment within the last ten years, is expected to be higher in the core (Bluestone, 1971:102–5) or in the state sector (O'Connor, 1973; Hodson, 1978:438–9). (Some other approaches emphasize that instability is related to seasonal causes in the core as compared to worker-related causes in the periphery: see Oster, 1979.) The results show, however, that only the hypothesis for union membership is supported. Contrary to expectations, there is higher current unemployment found in the core sectors.[8]

Though there are no clear predictions concerning occupational prestige, on the basis of findings in earlier research it is expected to be higher in the core. The results show that only the two taxonomies by Beck, Horan, and Tolbert produce the expected results; the taxonomy by Bibb and Form also produces a significant difference, but in the opposite direction. Since Bibb and Form constructed their taxonomy to examine differences between blue-collar workers in the two sectors, differences in overall occupational prestige is not of central importance to their model.

Labor Force Characteristics. One expects to find a more experienced work force in the core (see results obtained by Bluestone et al., 1973 in contrast to Averitt, 1968:128, on the concentration of elderly in the periphery). Similarly, a more educated and credentialed work force is expected to be found in the core. Degree attainment and the alternate measure, level of credential, are linked to educational certification arguments (Collins, 1979); years of schooling, which regulates workers' access to job ladders, is linked to internal labor market arguments (Doeringer and Piore, 1971, as borrowed by Beck et al., 1978:707).

The dual labor market prediction that blacks and women will be concentrated in the secondary labor market (Gordon, 1972:48; Reich et al., 1973:360) has been imported directly into the dual economy

[7] In contrast, Wallace and Kalleberg (1981:9) argue that unionization is most closely related to occupation, not industry. But occupation and structural elements in general—other than industry—have received little attention (for exceptions see Talbert and Bose, 1977; Stolzenberg, 1975; Pfeffer, 1977).

[8] Significant problems with these unemployment/stability of work variables have been noted elsewhere (see Beck et al., 1978:711, footnote 12). However, these findings were also replicated with the five years of the GSS, lending less credibility to the seasonal unemployment argument.

Table 1. Comparison of Mean Characteristics of Workers Using Different Sector Taxonomies

	Beck-Horan-Tolbert		Bibb-Form[a]		Hodson[b]			Tolbert-Horan-Beck	
	Core N = 1072	Periphery N = 512	Core N = 626	Periphery N = 908	Core N = 458	Periphery N = 989	State N = 137	Core N = 776	Periphery N = 808
DEPENDENT VARIABLES:									
Ln annual earnings[c]	9.0409	8.5110**	9.1496	8.6794**	9.1513	8.6983	9.1546*	9.1462	8.6040**
	(.8848)	(1.0720)	(.8147)	(1.0283)	(.8195)	(1.0317)	(.8057)	(.8150)	(1.0512)
Proportion poverty	.1147	.2695**	.0863	.2137**	.0852	.2144	.0730*	.0812	.2450**
workers[c]	(.3189)	(.4442)	(.2810)	(.4101)	(.2794)	(.4106)	(.2611)	(.2733)	(.4304)
LABOR MARKET:									
Proportion belonging to union[d]	.3022	.1445**	.3962	.1619**	.4148	.1628	.3431*	.3363	.1696**
Proportion unemployed at least once in past 10 years[e]	.3218	.3223	.3514	.3095**	.3712	.3084	.2555	.3466	.2983**
Proportion currently unemployed[f]	.0681	.0508*	.0958	.0407**	.1135	.0465	.0073	.0876	.0384**
Occupational prestige[g]	42.52	33.54**	38.51	40.44**	37.44	40.07	43.62	40.57	38.71**
LABOR FORCE:									
Work Experience									
Age for males[h]	39.94	39.29	39.62	39.48	39.83	40.04	38.22	39.35	40.37
Age for females[h]	38.25	39.30	38.26	38.56	38.68	38.89	34.53	37.67	39.11
Education									
Years of schooling[i]	12.87	11.51**	12.01	12.80**	11.88	12.60	13.01	12.41	12.45
Highest degree									
Proportion with less than high school[j]	.2127	.3496**	.2764	.2357**	.2948	.2568	.1314	.2294	.2834**
Proportion with high school or jr. college[j]	.5504	.5703	.5974	.5308**	.5873	.5258	.6788	.6005	.5149**
Proportion with B.A.[j]	.1604	.0723**	.1038	.1542**	.0939	.1446	.1679	.1340	.1300
Proportion with post-graduate degree[j]	.0765	.0078**	.0224	.0793**	.0240	.0728	.0219	.0361	.0718**
Level of credential[k]	2.1007	1.7383**	1.8722	2.0771**	k	2.0334	k	1.9768	1.9901
Sex/Race									
Proportion female[l]	.3610	.4668**	.1949	.5507**	.1921	.5137	.2190*	.2564	.5285**
Proportion nonwhite[m]	.0858	.1211**	.0942	.1024	.0895	.1031	.0803	.0786	.1151**
Parental Background									
Years parental schooling[n]	9.60	8.92**	8.98	9.71**	9.05	9.57	9.15	9.29	9.47

* p < .10; ** p < .05.

NOTE: Values based on deletion of 99 cases with missing income resulting from occupational changes (see Hauser, 1980). Values above also differ from those reported by Hauser because of differences in assignment of missing values and exclusion of cases (see below). Significance levels reported in this table indicate differences between means. Standard deviations are in parentheses.

[a] Sample size is smaller due to excluded industry group. See Appendix.

[b] The three sectors identified by Hodson are used in the regression analyses predicting the log of income reported in Tables 1 and 3. It was necessary to combine core and state in the logist predictions of proportion of poverty workers because of missing marginals primarily in the state sector for race, current unemployment, age for males, and proportion B.A. and post-graduate degrees. The following are the mean values (N = 595) for combined core and state sector (replaces values reported above the core sector): 9.1522**; .0824**; .3983**; .3446*; .0891**; 38.86**; 39.47; 37.63; 12.14**; .2571; .6084**; .1109**; .0235**; 1.9008**; .1983**; 0874; 9.07**. Significant differences are computed by comparing the means of the core-state to the periphery.

[c] See Beck et al. (1978:710) and Hauser (1980:705–6, Table 1, footnote 2).

[d] 1 = yes. Yes = NORC original codes 1 (R belongs) and 3 (both R and spouse); No = 2 (spouse), 8 (don't know) and 9 (no answer).

[e] 1 = yes. Missing data for seven cases are assigned the modal value, 1.

[f] 1 = unemployed.

[g] Missing data for 4 cases are assigned the modal value, 50.

[h] Missing data for 2 male ages and 3 female ages are assigned the modal value, 25.

[i] Missing data for 4 cases are assigned the modal value, 12.

[j] Dummy variable used in regression equations predicting income (Tables 1 and 3). Omitted category: proportion with less than high school. 1 = yes. Missing data assigned to modal category "high school or jr. college."

[k] Nominal variable used in logist analysis of proportion poverty workers. Values correspond to the four specific degree categories listed above.

[l] 1 = female.

[m] 1 = nonwhite.

[n] Parental education is taken as the years of schooling for the respondent's father. If this information is not available, data for mother's education or modal values were substituted.

literature (e.g., Bluestone et al., 1973:81–5; Hodson, 1978:434–6). Some recent evidence suggests that these groups, as expected, are concentrated in the periphery (Beck et al., 1978:710–1 and Tolbert et al., 1980:1112). This concentration is apparently independent of the major dimensions of dual economy at the institutional level, such as industrial concentration and assets (Oster, 1979).

There are no predictions about the distribution of parental background characteristics such as parental education, which is identified only as "characteristic of the individualist perspective" (Beck et al., 1978:708), with no reference to the dual economy or dual labor market literatures.

The results reported in Table 1 provide little support for the predictions concerning labor force characteristics. Expectations for work experience and education variables are generally disconfirmed. Not only are the differences not always statistically significant, but, more importantly, the differences are often in the wrong direction for the education variables. For example, contrary to expectation, significantly higher years of schooling and level of credential are found in the peripheral sector in the Bibb-Form taxonomies. The same was found for parental years of education. Sex is the only labor force characteristic which clearly confirms expectations, with women consistently concentrated in the periphery. While only significant in two of the taxonomies, the distribution of nonwhites is also as expected.

SECTORAL EARNINGS DETERMINATION

Differences in sector definition can alter the distribution of labor-market conditions and worker characteristics dramatically, and predictions concerning sectoral distributions are often not supported, with the direction of differences sometimes contrary to those predicted. However, it may still be the case that there are similar sectoral differences across all taxonomies in terms of the processes by which economic outcomes are determined. To examine these processes, two types of results are reported for each of the four

taxonomies: (1) Results from standard regressions (for predicting log of income) and from logistic regressions (for predicting proportion of workers with poverty earnings) which provide information on the sector-specific significance of individual variables, and (2) The analysis of sectoral interaction, using the Chow test or the likelihood-ratio test, as appropriate, to determine significant sectoral differences.[9]

Annual Earnings

Table 2, Panel A, presents the regression coefficients for the log of annual earnings and the significance of variables within each sector for all four taxonomies. Overall, a significant amount of variance is explained within each sector by the variables entered into the regression equations for all four taxonomies, though the individual variables found significant differ.

Labor Market Conditions. Examining the degree of significance of variables within the equation for each sector, labor market variables are found to be significant in all taxonomies. Union membership, work stability, and occupational prestige have positive effects, while unemployment has negative effects on income, with the exception of the Hodson state sector.

Labor Force Characteristics. Labor force variables show a much less consistent pattern across variables, with less consistency among the four core sectors than among the four peripheral sectors across taxonomies. Nonetheless, a few consistent patterns emerge. For example, the findings for each taxonomy show a significant negative effect of sex upon income,[10] as predicted, while the nonsignificance of the effects of race and years of education are unexpected findings. The contrasting effects of race and sex are particularly striking, in view of the tendency to group these variables together (for an example, see Oster, 1979; for an exception, see Bibb and Form, 1977, where sex alone is considered). Years of parental schooling has a consistent, surprising, negative impact on earnings in the periphery. The apparent anomalies and inconsistencies once again raise questions about the relationship between the underlying causes of industrial differentiation and the labor market and labor force consequences.

Sectoral Differences. In Table 2, the Pooled Equation with Sectoral Interactions reports sectoral differences. Only the Beck-Horan-Tolbert taxonomy produces significant differences, using the Chow test (as in Beck et al., 1978) and Fischer's protected t (Cohen and Cohen, 1975:162–7), when all of the interaction variables are added to the equation. Significance of individual variables is discussed below only when the overall Chow test results are significant. The results for individual variables should be taken as suggestive when the overall significance is not obtained, and hence are not discussed further here.[11] This result calls into question the validity of sector differentiation as currently operationalized in the other three taxonomies. Even for the Beck-Horan-Tolbert taxonomy, only three of the thirteen independent variables have

[9] Serious estimation problems are encountered when using standard regression techniques for dichotomous dependent variables (see Theil, 1971; Nerlove and Press, 1973). To overcome these problems, the logistic multiple regression model fitted to a single binary dependent variable is used (described fully in Harrell, 1980; see also Hanushek and Jackson, 1976). The logistic model is suited for use with continuous as well as polytomous independent variables, and hence more appropriate for analysis of the data used here than other available techniques. The likelihood ratio test (Maddala, 1977:43–5) is used to determine which taxonomies produce overall significant differences in terms of proportion with poverty earnings. For the standard multiple regression equations, the Chow test is used both to determine which taxonomies produce overall significant differences between sectors in terms of income and to determine if the criteria for Fisher's protected t are met (Cohen and Cohen, 1975:162–7).

[10] If hours or weeks worked had been controlled, the effect of sex would be reduced. However, these variables were not included in order to replicate the original work of Beck, Horan, and Tolbert (1978).

[11] For Bibb-Form, only occupational prestige and years of parental education emerged as significant. For Hodson, the following were significant, often marginally: current unemployment and age for males in both core and periphery; prestige, race, and parental education in the periphery only; high school/junior college in the core only. For Tolbert-Horan-Beck, prestige, postgraduate education, race, and years of parental education were significant.

Table 2. Comparison of Metric Regression Coefficients for Different Sectoral Taxonomies for Dependent Variable Income

	Beck-Horan-Tolbert[a]		Bibb-Form		Hodson[b]			Tolbert-Horan-Beck	
	Core	Periphery	Core	Periphery	Core	Periphery	State	Core	Periphery
A. Sector Specific									
Intercept	7.7620	7.4082	7.4469	7.5422	7.4064	7.6732	6.7596	7.5165	7.5470
LABOR MARKET:									
Union member (1 = yes)	.3427††	.3768††	.3055††	.3068††	.2942††	.3612††	.3128†	.2903††	.3447††
Work stability (1 = stable)	.2727††	.3084††	.2524††	.3163††	.3288††	.2886††	.1463	.2835††	.2916††
Unemployed (1 = yes)	−.2433†	−.2405*	−.2264**	−.3021**	−.2444**	−.2521**	.8049	−.2589†	−.2429*
Occupational prestige	.0112††	.0246††	.0122††	.0195††	.0119††	.0190††	.0140†	.0127††	.0178††
LABOR FORCE:									
Work Experience									
Age for males	.0121††	.0060*	.0133††	.0086†	.0096†	.0090†	.0253††	.0133††	.0087†
Age for females	.0075†	.0034	.0074*	.0046*	.0060	.0046**	.0062	.0079**	.0045*
Education									
Years of schooling	.0002	.0324	.0286*	.0088	.0557†	−.0057	.0009	.0279*	.0045
Highest degree									
High school/Junior college	.2900††	.2502**	.1345*	.2987†	.0132	.3617††	.5655†	.1675**	.3260†
B.A.	.4969††	.2436	.2531*	4509†	.1813	.5605††	.6238**	.2566**	.5045†
Postgraduate degree	.6787††	1.1402**	.4784**	.6665†	.2145	.8130††	.9860**	.3850**	.8221††
Sex/Race									
Sex (1 = female)	−.4765†	−.6401†	−.3579**	−.5048†	−.3805*	−.5101†	.0102	−.3998†	−.4960†
Race (1 = nonwhite)	−.0362	.0764	−.0843	.0296	−.2262**	.0733	.2177	−.0989	.1186
Parental Background									
Years parental schooling	−.0000	−.0463††	.0118*	−.0286†	.0050	−.0265††	.0238*	.0004	−.0258†
Standard error of estimate	.7242	.8639	.6735	.8148	.6778	.8258	.6047	.6594	.8564
R^2	.3382	.3671	.3309	.3812	.3354	.3677	.4906	.3564	.3472
F-ratio	41.59†	22.22†	23.28†	42.37†	17.24†	43.62†	9.11†	32.45†	32.49†
B. Pooled Equation with Sectoral Interactions									
F-ratio[c]	2.43**		1.40		1.43			.99	

* $p < .10$; ** $p < .05$; † $p < .01$; †† $p < .001$.

[a] Compare to Table 3 in Beck et al. (1978) and Table 3 in Hauser (1980). Refer to footnotes in Table 1 in present article for definition of sample and variables.

[b] In panel B, the regression equation with interaction variables includes the following dummy variables: (1) if in Hodson core sector, HC = 1 (HC = 0 if in periphery or state sectors), and (2) if in Hodson periphery sector, HP = 1 (HP = 0 if in core or state sectors). Thus, the Hodson state sector was the omitted category and always equal to zero.

[c] Chow test for the null hypothesis for no sectoral interactions, df = 13, 1557 for Beck et al. and Tolbert et al., df = 13, 1507 for Bibb-Form, and df = 26, 1543 for Hodson. There are three significant sector-variable interactions for Beck et al.: occupational prestige $p < .005$; age for males $p < .10$; and parental education $p < .0005$.

significantly different effects between sectors (see footnote c of Table 2).

Labor Market Conditions. Only the effects of occupational prestige on annual earnings differentiates sectors significantly. Union membership, work stability, and unemployment do not show significant differences. Since the only significant difference (occupational prestige) is one ambiguously predicted (and generally not even mentioned) in dual economy theory, interpretation is not clear. Certainly, overall support for dual economy predictions about labor market conditions is weak.

Labor Force Characteristics. Work experience for males has a marginally significant differential influence on income attainment between sectors: parental background has a much stronger effect. While work experience has some limited importance in dual economy theory, parental background does not. Again, the most critical variables—notably education and credentialing—fail to exhibit the expected differential influence.

Poverty Earnings

Table 3 presents the logist coefficients for poverty earnings (proportion of workers with poverty-level earnings) and the significance of variables within each sector. The D statistic measures the fit of the model, equivalent to "R square in the normal setting" (Harrell, 1980:83). Because the logist regression procedures require that there be marginals for each combination of variables, two changes in variable definition had to be made. First, Hodson's separate core and state sectors had to be combined: since the results obtained using the separate sectors are quite similar, such a combination appears justifiable (Wallace and Kalleberg, 1981, combine these sectors in all analyses).[12] Second, a new variable, credential, was constructed. Since all degree levels are included, no information is lost (see Table 1, footnote k).

Once again, a significant amount of the variance is explained for all four taxonomies in each sector. However, the pattern of significance of variables within each sector is less consistent across taxonomies and provides less support for dual economy predictions.

Labor Market Conditions. Here, three of the labor market variables are consistently significant across all taxonomies for all sectors, with union membership, work stability, and occupational prestige significantly decreasing the probability of poverty earnings. Unemployment uniformly increases the probability of poverty earnings, but the pattern of significance varies: unemployment has significant effects in the core for all taxonomies except Hodson, and significant effects in the periphery for Bibb-Form and Hodson.

Labor Force Characteristics. Again, as in the case of the annual income regressions, labor force characteristics show a much less consistent pattern across taxonomies than do labor market conditions. No variable is consistently significant in determining the probability of poverty earnings across both sectors in all taxonomies, though sex and years of parental schooling are significant in the periphery for all four. Interpretation of the effects of years of parental schooling is, however, difficult, since in all four taxonomies it increases the probabilty of poverty earnings in the periphery. Level of credentials operates as expected; in all the taxonomies, except Beck-Horan-Tolbert, it is significant in the periphery; in that one, it is significant in the core. What is notable is the general lack of significance of most labor force variables.

Sectoral Differences. Significant sectoral differences for the dichotomous dependent variable, proportion with poverty earnings, are examined using the likelihood ratio test (Maddala, 1977:43–5). The

[12] Two problems arise because of the smaller sample size in the public sector. First, there is a reduced likelihood of finding significant differences between means of characteristics comparing the three sectors. Nonetheless, we have used Hodson's trichotomous division, though interpretation of the results should be made cautiously. Second, it was impossible to estimate MLEs using logist regression procedures for the dependent variable, proportion of workers with poverty level income, if all three sectors were examined. Thus, Hodson's core and public sectors had to be combined.

Table 3. Comparison of Logist Regression Coefficients for Different Sectoral Taxonomies for Dependent Variable Proportion of Workers with Poverty-Level Earnings

	Beck-Horan-Tolbert[a]		Bibb-Form		Hodson		Tolbert-Horan-Beck	
	Core	Periphery	Core	Periphery	Core-State[b]	Periphery	Core	Periphery
Intercept	.8225	.1688	2.5474	.4369	2.4666	.3905	2.1546	.3572
LABOR MARKET:								
Union member (1 = yes)	−1.3527††	−1.4918††	−1.3689††	−1.2596††	−1.2524†	−1.4058††	−1.3160††	−1.3107††
Work stability (1 = stable)	−.6086†	−.8565††	−1.1643†	−.6100†	−1.1699†	−.6128†	−.9370†	−.6908††
Unemployed (1 = yes)	.9790†	.3904	.7672*	.8507*	.6490	.7875**	.9496**	.6504
Occupational prestige	−.0396††	−.0590††	−.0372**	−.0583††	−.0391**	−.0525††	−.0458†	−.0468††
LABOR FORCE:								
Work Experience								
Age for males	−.0278*	.0152	−.0236	−.0048	−.0230	−.0025	−.0346*	.0025
Age for females	−.0029	−.0050	−.0103	−.0070	−.0266	−.0026	−.0287	−.0006
Education								
Years of schooling	.0609	−.0188	−.1184	.0186	−.0461	.0232	−.0337	.0079
Highest Degree								
Credential	−.7256**	−.4039	−.3453	−.4515*	−.6271	−.5610**	−.5344	−.5114**
Sex/Race								
Sex (1 = female)	.6369	2.0314†	.8968	1.4524**	1.4726	1.3085**	1.1725	1.3812**
Race (1 = nonwhite)	−.3373	−.3380	−.9566	−.0786	−.8256	−.2459	−.8300	−.3153
Parental Background								
Years parental schooling	.0218	.1158†	−.0254	.0864†	−.0520	.0892†	−.0023	.0850†
D	.125††	.192††	.120††	.181††	.113††	.175††	.111††	.176††
−2 Log L	612.73	477.63	284.46	744.13	263.98	821.00	341.36	729.67

* p < .10; ** p < .05; † p < .01; †† p < .001.

NOTE: Likelihood ratio test results for all four taxonomies show no significant interactions between the sectoral variable and the independent variables.

[a] Compare to Table 3 in Beck et al. (1978). Refer to footnotes in Table 1 for definition of sample and variables.

[b] Core and state sectors had to be combined because the small sample size of the state sector resulted in missing marginals for some combinations of variables which prevented logist estimation.

results provide strong negative evidence for dual economy predictions: none of the taxonomies have significant interactions between the sectoral variable and the independent variables. Parental education was significant in all four taxonomies; age for males was significant in Beck-Horan-Tolbert, but since results concerning the significance of individual variables are suspect when overall significance is not obtained, they will not be discussed further.

CONCLUSIONS AND IMPLICATIONS

On the whole, the results provide little evidence of sectoral differentiation. At the same time, the use of labor market and labor force variables selected by Beck, Horan, and Tolbert yield good explanations of income and proportion with poverty earnings within each sector. However, the patterns of significance for individual variables are frequently inconsistent across taxonomies, and the predictions derived from dual economy theory are often not supported. The results for labor market variables are more often supportive than the results for the variables tapping labor force characteristics. Finally, the crucial tests for significant sectoral differences demonstrate that, with one exception, income and proportion with poverty workers are determined similarly across sectors.

These findings indicate that current work on dual economy needs to be reconsidered, and that some reformulation is necessary. Taxonomies are frequently used uncritically, as if they in fact tapped significant institutional differences (Beck et al., 1980; Otto et al., 1980). Agreement on a single taxonomy might lead to more systematic refinement of the dual economy approach, but there are several obstacles to developing such a taxonomy: (1) the differing emphases that theoretical traditions place on institutional aspects of the economy, e.g., importance of the public or state sector; (2) lack of consensus concerning causal relations among economic scale variables, labor-market conditions, and worker characteristics; (3) other sources of structural heterogeneity within sectors, e.g., occupations which differ in wage rates (Stolzenberg, 1975; Freedman, 1976; Rees and Schultz, 1970) and distribution (Wallace and Kalleberg, 1981) across sectors; (4) lack of detail in industrial classifications (using two-digit SIC classification), artificially creating differences in assignment to sector, with some of the combined industries belonging in the core, others in the periphery (see Tolbert et al., 1980).

The argument that the institutional structure of the economy is important in determining economic processes and outcomes is not disputed. On the whole, the proposed taxonomies have served as heuristic devices in the early stages of investigation of dual economy theory. However, the divergent findings which emerge using these four taxonomies demonstrate the need for further comparative examination and clear theoretical specification. Empirical derivation of taxonomies seems unlikely to provide a satisfactory solution, since different investigators report strikingly different relations between key variables. The need for further development of the underlying theoretical models is clear.

APPENDIX SECTORAL CLASSIFICATIONS

Industry Group	Beck, Horan, and Tolbert (1978)	Bibb-Form (1977)	Hodson (1977)	Tolbert, Horan, and Beck (1980)
Agriculture, forestry, and fisheries	Periphery	Excluded	Periphery	Periphery
Mining				
Metal mining	Core	Core	Core	Core
Coal mining	Core	Core	Core	Core
Crude petroleum and natural gas	Core	Core	Core	Core
Nonmetalic mining and quarrying	Core	Core	Periphery	Core
Construction	Core	Core	Core	Core

APPENDIX Continued

Industry Group	Beck, Horan, and Tolbert (1978)	Bibb-Form (1977)	Hodson (1977)	Tolbert, Horan, and Beck (1980)
Durable Manufacturing				
Lumber and wood products	Periphery	Core	Periphery	Periphery
Furniture and Fixtures	Periphery	Core	Periphery	Periphery
Stone, clay, and glass products	Core	Core	Core	Core
Metal industries	Core	Core	Core	Core
Machinery, except electrical	Core	Core	Core	Core
Electrical machinery, equipment, supplies	Core	Core	Core	Core
Transportation equipment	Core	Core	Core	Core
Professional and photographic equipment	Core	Core	Core	Core
Ordinance	Core	Core	Core	Core
Miscellaneous durable manufacturing	Periphery	Core	Core	Periphery
Nondurable Manufacturing				
Food and kindred products	Periphery	Periphery	Core	Core
Tobacco manufacturers	Periphery	Periphery	Core	Core
Textile mill products	Periphery	Periphery	Periphery	
Knitting				Periphery
Dyeing and finishing				Core
Textile floor coverings				Periphery
Yarn, thread, and fabric mills				Core
Miscellaneous textile mills				Periphery
Apparel and other fabricated textiles	Periphery	Periphery	Periphery	Periphery
Paper and allied products	Core	Periphery	Core	Core
Printing, publishing, and allied industries	Core	Periphery	Periphery	Core
Chemicals and allied products	Core	Core	Core	Core
Petroleum and coal products	Core	Core	Core	Core
Rubber and miscellaneous plastic products	Core	Core	Core	
Rubber products				Core
Miscellaneous plastic products				Periphery
Leather and leather products	Periphery	Periphery	Periphery	
Tanned, finished				Periphery
Footwear				Core
Leather products, except footwear				Periphery
Not specified nondurable manufacturing	Periphery	Periphery	Periphery	Periphery
Transportation				
Railroads and railway express service	Core	Core	Core	Core
Street railways and bus lines	Core	Core	Periphery	Periphery
Taxicab service	Core	Core	Periphery	Periphery
Trucking service	Core	Core	Periphery	Core
Warehousing and storage	Core	Core	Periphery	Core
Water transportation	Core	Core	Periphery	Core
Air transportation	Core	Core	Core	Core
Petroleum and gasoline pipelines	Core	Core	Periphery	Core
Services incidental to transportation	Core	Core	Periphery	Periphery
Communications				
Radio broadcasting and television	Core	Core	Periphery	Core
Telephone (wire and radio)	Core	Core	Core	Core
Telegraph (wire and radio)	Core	Core	Core	Core
Utilities and sanitary services				
Electric light and power	Core	Core	State	Core
Gas, steam, and supply systems	Core	Core	State	Core
Electric-gas utilities	Core	Core	State	Core
Water supply	Core	Core	Periphery	Periphery
Sanitation services	Core	Core	Periphery	Periphery
Other not specified utilities	Core	Core	Periphery	Periphery

APPENDIX Continued

Industry Group	Beck, Horan, and Tolbert (1978)	Bibb-Form (1977)	Hodson (1977)	Tolbert, Horan, and Beck (1980)
Wholesale trade	Core	Periphery	Periphery	
Motor vehicles and equipment				Periphery
Drugs, chemicals				Core
Dry goods and apparel				Periphery
Food and related products				Core
Farm products				Periphery
Electrical goods				Core
Hardware, etc.				Periphery
Machinery and equipment				Core
Metals and minerals				Core
Petroleum products				Periphery
Scraps and waste materials				Periphery
Alcoholic beverages				Core
Paper and its products				Periphery
Lumber and construction				Periphery
Wholesalers n.e.c.				Periphery
Not specified				Periphery
Retail trade	Periphery	Periphery	Periphery	Periphery
Finance and insurance	Core	Periphery	Periphery	Core
Real estate	Core	Periphery	Periphery	Periphery
Business and repair services	Periphery	Periphery	Periphery	Periphery
Personal services	Periphery	Periphery	Periphery	Periphery
Entertainment and recreation services	Periphery	Periphery	Periphery	Periphery
Professional and related services	Core	Periphery	Periphery	
Offices of physicians/dentists				Core
Hospitals/convalescent institutions				Periphery
Legal services				Core
Educational services				Periphery
Museums, etc.				Periphery
Religious organizations				Periphery
Welfare services				Periphery
Nonprofit organizations				Periphery
Engineering and architectural				Core
Accounting, auditing, bookkeeping				Core
Miscellaneous professional				Core
Public administration	Core	Core	State	Core

REFERENCES

Averitt, Robert T.
1968 The Dual Economy: The Dynamics of American Industry Structure. New York: Horton.

Baron, James N. and William T. Bielby
1980 "Bringing the firms back in: stratification, segmentation, and the organization of work." American Sociological Review 45:737–65.

Beck, E. M., Patrick M. Horan, and Charles M. Tolbert II
1978 "Stratification in a dual economy: a sectoral model of earnings determination." American Sociological Review 43:704–20.
1980 "Industrial segmentation and labor market discrimination." Social Problems 28:113–30.

Bibb, Robert and William H. Form
1977 "The effects of industrial, occupational, and sex stratification on wages in blue-collar markets." Social Forces 55:974–96.

Blau, Peter M. and Otis Dudley Duncan
1967 The American Occupational Structure. New York: Wiley.

Bluestone, Barry
1971 "The characteristics of marginal industries." Pp. 102–7 in David M. Gordon (ed.), Problems in Political Economy: An Urban Perspective. Lexington, Mass.: D. C. Heath.

Bluestone, Barry, William M. Murphy, and Mary Stevenson
1973 Low Wages and the Working Poor. Ann Arbor, Michigan: Institute of Labor and Industrial Relations.

Cain, Glenn G.
1976 "The challenge of segmented labor market theories to orthodox theory: a survey." Journal of Economic Literature 14:1215–57.

Cohen, Jacob and Patricia Cohen
1975 Applied Multiple Regression: Correlation Analysis for the Behavioral Sciences. Somerset, N.J.: Wiley.

Collins, Randall
1979 The Credential Society. New York: Academic Press.
Davis, James A.
1975 National Data Program for the Social Sciences: Codebook for the Spring 1975 General Social Survey. Chicago: National Opinion Research Center.
1976 National Data Program for the Social Sciences: Codebook for the Spring 1976 General Social Survey. Chicago: National Opinion Research Center.
Doeringer, Peter B. and Michael J. Piore
1971 Internal Labor Markets and Manpower Analysis. Lexington, Mass.: Lexington.
1975 "Unemployment and the 'dual labor market'." The Public Interest 38:67–79.
Duncan, Otis Dudley, David L. Featherman, and Beverly Duncan
1972 Socioeconomic Background and Achievement. New York: Seminar Press.
Edwards, Richard C.
1975 "The social relations of production in the firm and labor market structure." Pp. 3–26 in R. C. Edwards, M. Reich, and D. M. Gordon (eds.), Labor Market Segmentation. Lexington, Mass.: D. C. Heath.
1979 Contested Terrain: The Transformation of the Workplace in the Twentieth Century. New York: Basic Books.
Freedman, Marcia K.
1976 Labor Markets: Segments and Shelters. Montclair, New Jersey: Allanheld, Osmun.
Freiman, Marc P.
1976 "Empirical tests of dual labor market theory and hedonic measures of occupational attainment." Ph.D. dissertation. Department of Sociology, Madison: University of Wisconsin.
Gordon, David M.
1972 Theories of Poverty and Underemployment. Lexington, Mass.: Heath.
Hanushek, Eric and John Jackson
1976 Statistical Methods for Social Scientists. New York: Academic Press.
Harrell, Frank
1980 "The LOGIST Procedure." Pp. 83–102 in SAS Supplemental Library User's Gude, 1980 Edition. Cary, N.C.: SAS Institute.
Hauser, Robert M.
1980 "On 'stratification' in a dual economy." American Sociological Review 45:702–12.
Hodson, Randy D.
1977 "Labor force participation and earnings in the core, peripheral, and state sectors of production." M.A.Thesis, Department of Sociology, University of Wisconsin, Madison.
1978 "Labor in the monopoly, competitive, and state sectors of production." Politics and Society 8:429–80.
Hodson, Randy D. and Robert L. Kaufamn
1981 "Circularity in the dual economy: a comment on Tolbert, Horan, and Beck." American Journal of Sociology 86:881–7.
Horan, Patrick M., Charles M. Tolbert, II, and E. M. Beck
1981 "The circle has no close." American Journal of Sociology 86:887–94.
Kalleberg, Arne L., Michael Wallace, and Robert P. Althauser
1980 "Economic segmentation, worker power, and income inequality." Unpublished manuscript, Indiana University.
Kaufman, Robert L., Randy Hodson, and Neil D. Fligstein
1981 "Defrocking dualism: a new approach to defining industrial sectors." Social Science Research, forthcoming.
Maddala, G. S.
1977 Econometrics. New York: McGraw-Hill.
Nerlove, Marc and S. James Press
1973 Univariate and Multivariate Log-Linear and Logistic Models. Santa Monica, Ca.: RAND.
O'Connor, James
1973 The Fiscal Crisis of the State. New York: St. Martins.
Oster, Gerry
1979 "A factor analytic test of the theory of dual economy." Review of Economics and Statistics 61:33–9.
Osterman, Paul
1975 "An empirical study of labor market segmentation." Industrial Labor Relations Review 28:508–23.
Otto, Luther B., Kenneth I. Spenner, and Vaughn R. A. Call
1980 "Careers and career lines: concepts, measures and explanations." Mimeographed paper. Boys Town, Neb.: Boys Town Center for the Study of Youth Development.
Pfeffer, Jeffrey
1977 "Toward an examination of stratification in organizations." Administrative Science Quarterly 22:553–67.
Pfeffer, Jeffrey and Jerry Ross
1979 "Do wage and status attainment differ by industry sector?" Unpublished manuscript, Stanford University.
Piore, Michael J.
1973 "Fragments of a sociological theory of wages." American Economic Review 63:377–84.
1975 "Notes for a theory of labor market stratification." Pp. 125–50 in R. C. Edwards, M. Reich, and D. M. Gordon (eds.), Labor Market Segmentation. Lexington: D. C. Heath.
Rees, Albert and George P. Shultz
1970 Workers and Wages in an Urban Labor Market. Chicago: University of Chicago.
Reich, Michael, David M. Gordon, and Richard C. Edwards
1973 "Dual labor markets: a theory of labor market segmentation." American Economic Review 63:359–65.
Rosenberg, Samuel
1976 "An empirical test of the dual labor market hypothesis." Unpublished manuscript, Williams College.
Sewell, William H. and Robert M. Hauser (eds.)
1975 Education, Occupation, and Earnings:

Achievement in the Early Career. New York: Academic Press.

Stolzenberg, Ross
1975 "Occupations, labor markets and the process of wage attainment." American Sociological Review 40:645–65.

Talbert, Joan and Christine E. Bose
1977 "Wage-attainment processes: the retail clerk case." American Journal of Sociology 83:403–24.

Theil, Henri
1971 Principles of Econometrics. New York: Wiley.

Tolbert, Charles M., II, Patrick M. Horan, and E. M. Beck
1980 "The structure of economic segmentation: a dual economy approach." American Journal of Sociology 85:1095–116.

Vietorisz, Thomas and Bennett Harrison
1973 "Labor market segmentation: positive feedback and divergent development." The American Economic Review 63:366–76.

Wallace, Michael and Arne L. Kalleberg
1981 "Economic organization, occupations, and labor force consequences: toward a specification of dual economy theory." To be published in Ivar Berg (ed.), Sociological Perspectives on Labor Markets. New York: Academic Press.

Zucker, Lynne G.
1979 "Sources of change in institutionalized systems: effects on occupational mobility." Revised version of a paper presented at the annual meeting of the American Sociological Association, September 1977.

STRUCTURAL CONSTRAINTS AND INDIVIDUAL CAREER EARNINGS PATTERNS*

ROBERT NASH PARKER
The University of Akron

The impact of recurrent economic cycles of recession and boom on individual earnings over time has been neglected in sociological research. Five hypotheses are derived which predict relationships between individual earnings and periodic economic fluctuations, as measured by the unemployment rate, for workers in general and for workers in certain social structural locations, defined by education, economic sector, union membership, and race. Results from spectral analyses indicate that periodic economic fluctuations constrain earnings in general over the short term—that is six to seven months. The importance of social structural location is underscored by the finding that those in more advantaged structural locations have earnings patterns which are not related to changes in the unemployment rate, while those in less advantaged structural locations have earnings patterns which are strongly related to changes in the unemployment rate. Both black and white earnings patterns are related to periodic economic fluctuations.

INTRODUCTION

"Recession," "inflation," "boom and bust," "stagflation," and "depression" have become part of the vernacular in modern, industrialized market economies. Despite the importance of economic fluctuations, since World War II, little sociological research has explored their relationships to individual earnings, employment, and occupational attainment. This paper focuses on the effects of periodic economic fluctuations on individual earnings patterns as further affected by individuals' social structural positions. These issues are investigated in terms of the expectations of sociological and macroeconomic theory.

* Direct all correspondence to Robert Nash Parker, Department of Sociology, The University of Akron, Akron, Ohio 44325.

This is a revised version of a paper presented at the annual meetings of the American Sociological Association, New York, 1980. I would like to thank William Form, editor, and two anonymous reviewers of ASR for suggestions which substantially improved this paper. I would also like to thank Alan C. Kerckhoff and Richard T. Campbell for helpful comments on earlier drafts and Elizabeth Mutran for technical assistance.

Economic and Sociological Approaches

Aggregate macroeconomic theory (e.g., Miller, 1973) provides some basic notions concerning the relationship between periodic economic fluctuations and individual earnings. In such an approach, the demand for workers varies inversely with the wage rate, so that, as wages increase, fewer workers are employed. However, as the wage rate increases, so does the available supply of workers. The intersection of the demand and supply curves, at the 'equilibrium' point, indicates the number of workers employed and the level of their wages. During a recession, this equilibrium will be upset by an aggregate decline in the demand for workers. This situation creates a temporary excess of workers seeking jobs (increased unemployment), which decreases the wage rate, since employers are able to offer lower wages. In periods of economic growth, an increase in the aggregate wage rate is predicted in the short run, because the increase in demand for workers causes a temporary shortage of workers seeking employment. In the moderate to long term, the supply of workers is elastic, so that equilibrium will be re-established by changes in the number of workers seeking employment in response to the aggregate demand and the wage rate. However, this readjustment takes time, because of both the duration of the fluctuation itself and the time necessary for information to be assimilated to the economy. In sum, macroeconomic theory states that if, in the short run, unemployment is high, individual earnings are lowered or at least increase less rapidly. Conversely, if unemployment is low, individual earnings rise or increase more rapidly.

This statement, of course, ignores much of the complexity of economic theory. More importantly, it violates a fundamental sociological position: that workers are differentiated by social backgrounds, resources, skills, and experiences concomitant with their location in the social structure. Early status attainment research focused on the manner in which social background influences educational attainment and the prestige of a worker's initial job (Blau and Duncan, 1967). Later research found that these structural characteristics influenced later occupational and earnings attainment (Sewell and Hauser, 1972:859). Education, including both academic and post-school training, has a central place in the human-capital perspective, which sees education as individual "investment" made to attain higher returns in prestige and earnings (Becker, 1964; Mincer, 1974). Thus, those with less education should have earnings which reflect the ups and downs of the economy more than the earnings of better educated coworkers.

While status-attainment and human-capital approaches have focused on individual characteristics as structural location indicators, others have argued that the structural characteristics of an individual's place of employment are important for attainment of earnings and prestige. For example, Bibb and Form (1977) and Beck et al. (1978) contrast explanations of social and economic attainment based on status attainment or human capital perspectives with those based on the structural characteristics of industries. The latter are classified as "core" industries (those with large amounts of capital investment, bureaucratization, and monopolistic tendencies) and "periphery" industries (those in more competitive markets, with lower levels of bureaucratization, profit margins, and capital investment) (see Bluestone et al., 1973). This "dual economy" approach implies that workers in peripheral industries, because of the aforementioned structural characteristics, should have earnings patterns which reflect the influence of recurring economic fluctuation; those in core industries should have patterns which are relatively protected from such fluctuations.[1]

Two other indicators of structural location also influence the earnings and career patterns of workers. Spilerman (1977:581) underscores the importance of unionization; union members should be more protected from unemployment and declining

[1] Although Hauser's (1980) criticism of Beck et al. (1978) may cast empirical doubt on dual economy research, it does not demonstrate major theoretical inadequacies.

earnings than nonunion members because unions are able to shield members from the impact of economic fluctuations via collective bargaining agreements concerning wages, hiring and firing priorities, and the distribution of available overtime opportunities. In addition, racial differences in earnings have been consistently documented by the status-attainment, human-capital, and dual-economy approaches (see Duncan, 1969). Thus, black workers' earning patterns should be more strongly related to periodic economic fluctuations than the patterns of white workers.

In summary, it is hypothesized that the earnings patterns of individual workers over time are constrained by the effects of periodic economic fluctuations, and that the degree of constraint varies directly with the social structural location of the individual.

DATA

The data to be analyzed are from the 1970 Detroit Area Study, "Careers in Detroit," conducted under the direction of Robert E. Cole (1977). The study obtained the responses of 638 male residents of Detroit on 633 variables focusing on job histories, career patterns, and job satisfaction. The study provides up to four sets of variables per respondent describing the entire work and earnings history. This study has the advantage of being entirely contained in one general labor market; factors such as industrial and racial-ethnic composition, transportation systems, and geographic location, which could affect employment variations, are controlled in the sense that they remain constant for all respondents.

On the basis of the four sets of job descriptions, 319 respondents were selected who had the most complete job histories.[2] Table 1 examines the four structural location variables plus average age for the 638 respondents in the total sample and the subset of 319 cases with the most complete job histories. Although breakdowns on race, education, and age are almost identical in the two groups, employment-sector and union-membership distributions are somewhat different. A respondent was considered in the primary sector if at least one-half or more of the jobs described by him were in the primary sector. Similarly, if the respondent was a union member in one-half or more of the jobs held during the period described by the four descriptor sets, he was classified as part of the union subgroup. A shift of approximately 10% from core to periphery and from no union membership to union membership occurs in the subsample. Since the selection was based on complete histories rather than stable or long-term histories, speculation about the effects of the selection on the earnings data is probably unnecessary. Since the magnitude of these shifts is not so great as to suggest their complete inacceptability, I conclude that the subsample represents workers in Detroit with the most complete work histories. This is not to say that the potentiality of bias due to selection has been eliminated, but that I cannot provide a reasonable description of any particular

Table 1. Comparison of Samples, Before and After Selection for Most Complete Work Histories

	Original Sample	Subsample
Sector of Employment		
Primary	428	182
	(67.2)	(57.1)
Secondary	209	137
	(32.8)	(42.9)
Union Membership		
Yes	296	182
	(47.7)	(57.4)
No	324	135
	(52.3)	(42.6)
Education		
HSG+	418	208
	(65.6)	(65.4)
<HSG	219	110
	(34.4)	(34.6)
Race		
Black	105	58
	(16.5)	(18.2)
White	533	261
	(83.5)	(81.8)
Average Age	38.75	39.56
Standard Deviation	11.42	10.58

NOTE: Figures in parentheses are percentages.

[2] In those cases in which fewer than four descriptive sets of variables were necessary to delineate a complete job history, duplicate values rather than missing value codes were recorded. Such cases were, therefore, included in the subsample.

bias or inconsistency which would seriously impair any conclusions to be drawn from the analysis below.

Construction of the Earnings Series

Several assumptions have been made in the derivation of the earnings life histories to be analyzed. The set of descriptor variables for each job history contained seven items from which individual earnings series were constructed. Three variables indicated the year at which the state of earnings began and ended and its duration in months, while four variables described hours worked and hourly wages earned at the beginning and end of the state. Earnings of salaried personnel were converted to hourly wages at the coding stage (Cole, 1977:53). Quarterly-earnings time series were calculated based on these seven pieces of information. The first assumption made in this process was that any change in earnings which occurred between the beginning and ending of each job was equally distributed over the duration of the job. Although this is an unrealistic model of earning change over time, information about the nature of such changes was not available. Alternatively, I might have assumed that earnings remained the same over the duration of the job but changed an amount equal to the difference between the beginning and ending quarterly wage in the last quarter the job was held. This is a most unrealistic model and would result in a series that appears to change more unevenly than is actually the case. A second operational assumption was that the average number of weeks worked during each year per respondent was equal to 48. Findings from the Parnes Survey (1970) suggest that this assumption may provide more underestimation than overestimation. In 1965, one of the years for which data is available in this study, slightly over 80% of males aged 45–59 in the survey worked 49 to 52 weeks during the year; only 10% of the sample worked less than 40 weeks.

The final step in the construction of the earnings series was the calculation of total subsample and subgroup means for each quarter covered by the data. Table 2 gives the number of cases available at the beginning of the period covered by the data and at the end of the period for the subsample and each subgroup. The closer to the earliest quarter in the series, ending in March, 1955, the fewer cases available for each mean, not because of missing data but rather because the respondents had differing amounts of labor market experience. In each group the number of cases available at the earliest point in the data is 15% or more of the number available at the latest point. Means were calculated for each group in each quarter in order to achieve a greater degree of stability than what might be the case if individual earnings series were analyzed.

Table 2. Number of Cases Available for Mean Earnings Series

Group	N of Cases, June 1970	N of Cases, March 1955
Subsample	316	80
Whites	259	66
Blacks	57	14
HSG+	206	41
<HSG	109	39
Primary sector	180	55
Secondary sector	136	25
Union	180	55
Nonunion	134	24

Unemployment rates, the indicator of periodic fluctuations, were available for Detroit from 1955 on. Bimonthly data from 1955 to 1961 (U.S. Department of Labor, 1955–1961) and from 1962 to 1970 (U.S. Department of Labor, 1962–1970) were used to calculate a quarterly unemployment series for the Detroit SMSA. In summary, the data to be analyzed consist of quarterly time series of the unemployment rate and mean quarterly earnings for the total subsample and eight subgroups for the Detroit SMSA covering the period March 1955 to June 1970.

METHODS

The time series described above are analyzed via spectral and cross-spectral techniques. The mathematical foundation of spectral analysis falls under the heading of Fourier analysis, the basic result of which is that any discrete stochastic or random function can be represented as the sum of a series of sines and cosines (Fishman, 1969:18; Mayer and Arney, 1974:316;

Bloomfield, 1976:1). The implication of this result is that any time series can be decomposed into the periodic components of varying durations which when summed result in the observed series. Any observed time series is the result of a number of distinct components that act simultaneously to produce the individual observations in the series. Especially when some of these components are cyclical in nature, traditional time series regression and correlation methods will obscure the impact of these components because in such approaches the emphasis is on the discovery and analysis of overall trends and/or relationships. Spectral analysis focuses on periodic components, and it provides estimates of the importance of each periodic component in relation to the sum of all periodic components. Cross-spectral analysis, in which the relationships among periodic components in two time series can be evaluated, is analogous to regression analysis. In addition to providing estimates of the strength of relationships between components of the same and differing durations across the two series given by the coherence squared (K^2),[3] the phase angle can be used to determine if the periodic component in the independent series leads or lags the periodic component in the dependent series that is related to it.

The major assumption required for spectral analysis is that the time series be covariance stationary, that is, have a constant mean over time. Although most time series usually exhibit strong first order serial correlation, and are therefore not covariance stationary, the application of a simple linear transformation will often result in the removal of such trends. The transformed series can then be tested for stationarity with a variant of the Kolmogorov-Smirnov test developed for this purpose by Bartlett (1966).

The spectrum of any series is likely to be radically affected by small random variations which can easily obscure the true important variance components (Mayer and Arney, 1974:326). It is standard practice to estimate the spectrum or cross-spectrum using a weighting function or spectral window. The resulting estimated variances at each frequency are a weighted average of the variance at the wavelength of interest and the variances at wavelengths of longer and shorter durations. This procedure smooths the spectrum so that it becomes more difficult for a spurious or random variance component to appear more important than other components (see Fishman, 1969:87–102). The spectra analyzed below were smoothed with the Parzen window (Fishman, 1969:101–2), with the number of lead and lag periods over which components were averaged equal to 16, and the number of quarters in each series equal to 61.

The analysis will proceed by first computing the spectra of the ten original series (total subsample, high and low education, core and periphery employment, union and nonunion membership, blacks and whites, and the unemployment series) and testing for stationarity. If stationarity is found to be absent, a first difference transformation (See Tintner, 1940) will be applied and the transformed spectra will again be tested for stationarity. If the transformed series are found to be stationary, or if the degree of nonstationarity has been substantially reduced,[4] the cross-spectra for unemployment and the nine earnings series will be calculated and five hypotheses will be evaluated via the significance of specific period components, the K^2s and associated phase angles.

A problem often raised in the literature dealing with cross-spectral analysis is that of alignment or time coordination (Mayer and Arney, 1974:325). Although misalignment can cause severe bias in the estimates of the cross-spectrum, the prob-

[3] The coherence squared or K^2 is a direct analog to R^2, as it measures the degree of association of the spectra of two or more series at identical frequencies. As with R^2, the evaluation of null hypotheses of the form $K^2 = 0$ can be performed (Fishman, 1969:137).

[4] Although it would be ideal to find all nonstationarity removed by the relatively simple transformation employed here, this is not always the case. However, Granger and Hatanaka (1964:145–6) argue that if *nonstationarity* is substantially reduced, it is reasonable to proceed with the analysis.

lem is particularly acute when a strong cross-lag relationship exists between two series at a distance greater than in the number of lags used in the spectral window (Bloomfield, 1976:229). In the particular case of the unemployment and earnings series considered herein, the strongest relationship exists at the instantaneous period.[1]

RESULTS AND DISCUSSION

Table 3 contains results from the stationarity tests for the nine earnings series and the unemployment series. In each case the original data exhibit significant nonstationarity. The application of the first difference transformation removes the nonstationarity in each of the earnings series quite successfully; and in the case of employment, this transformation substantially reduces the degree of nonstationarity.

Table 3. Tests for Stationarity, Original and First Variate Difference Filtered Series

Series	Chi-Square Values	
	Original Series	Filtered Series
Unemployment	27.1†	11.3**
Total mean earnings	31.0†	2.0
Whites	29.7†	2.1
Blacks	37.8†	4.8
HSG+	33.6†	0.5
<HSG	27.4†	5.4
Primary	36.5†	0.4
Secondary	18.3†	4.8
Union	41.8†	4.6
Nonunion	18.0†	3.1

NOTE: See Blalock (1972:264) for the computation of chi-square values from the Kolmogorov-Smirnov test statistic D; see Bartlett (1966) for the derivation of D.

* Significant at .05 level, 2df.
** Significant at .01 level, 2df.
† Significant at .001 level, 2df.

Tables 4 and 5 contain a summary of the results of the estimated cross-spectra for

Table 4. Cross-Spectra Results: Unemployment with Earnings

Unemployment series with	Periodic Components with Duration Measured in Quarters						
	2.44	2.35	2.26	2.18	2.10	2.03	2.00
Total subsample							
K^2	—	—	—	—	—	.238*	.242*
Phase						3.071	3.086
White							
K^2	.252*	.305*	.350**	.386**	.410†	.422†	.421†
Phase	2.582	2.677	2.751	2.810	2.861	2.908	2.954
Black							
K^2	—	—	—	.253*	.280*	.297*	.303*
Phase				3.054	3.058	3.063	3.069
HSG+							
K^2	—	—	—	—	—	—	—
Phase							
<HSG							
K^2	—	—	—	.235*	.255*	.267*	.270*
Phase				2.939	2.956	2.973	2.991
Primary sector							
K^2	—	—	—	—	—	—	—
Phase							
Secondary sector							
K^2	—	.240*	.281*	.312*	.333**	.344**	.344**
Phase		2.646	2.722	2.777	2.820	2.857	2.893
Union members							
K^2	—	—	—	—	—	—	—
Phase							
Nonunion							
K^2	—	—	—	—	.255*	.268*	.272*
Phase					3.037	3.059	3.079

— Insignificant.
* Significant at .20 level.
** Significant at .10 level.
† Significant at .05 level.

Table 5. Percent of Total Variance Accounted for by Periodic Components with Significant K^2: Unemployment with Earnings

Percent of Total	Duration Measured in Quarters						
Variance in	2.44	2.35	2.26	2.18	2.10	2.03	2.00
Total subsample	—	—	—	—	—	12.57	12.82
Whites	5.10	6.38	7.61	8.69	9.57	10.14	10.35
Blacks	—	—	—	11.41	12.60	13.40	13.68
HSG+	—	—	—	—	—	—	—
<HSG	—	—	—	11.09	11.19	11.89	12.14
Primary	—	—	—	—	—	—	—
Secondary	—	6.59	8.12	9.48	10.58	11.30	11.55
Union	—	—	—	—	—	—	—
Nonunion	—	—	—	—	11.81	12.52	12.76

— Insignificant K^2 (see Table 4).

the unemployment series and each of the nine earnings series. The numbers across the top of Table 4 correspond to periodic components with wavelengths equal to the time-domain duration indicated; that is, the first column of Table 4 corresponds to a cycle occurring in both the unemployment and earnings series which takes approximately 2.44 quarters or a little over 7 months to complete. As can be seen clearly from the durations listed in Table 4, these data yield absolutely no evidence of any long-term periodic components; significant relationships between unemployment and earnings were found only for periodic components with durations given across the top of Table 4.

Results in Table 4 for the entire subsample indicate the existence of short-term cycles in unemployment and earnings with durations of 2.03 and 2.00 quarters. The associated phase statistics indicate that periodic changes in unemployment occur first, and are then followed by similar short-term changes in earnings. Thus, the general notion that changes in unemployment rates are related to changes in individual earnings is supported. Table 5 further substantiates this hypothesis, since these two periodic components alone account for over 25% of the joint variance in unemployment rates and individual quarterly earnings.

Hypotheses predicting structural differences in this general relationship are strongly supported by the results in Tables 4 and 5 in three of the four comparisons. Those with a completed high-school education or greater, those who work in core industries, and those who were union members were completely protected from periodic fluctuations as measured by unemployment rates; for these three groups there were no significant relationships between their earnings and unemployment rates at any frequency or duration. However, earnings of those with lower levels of educational attainment, in periphery industries, and without union memberships were significantly associated with short-term variations in unemployment rates. Table 5 indicates that substantial proportions of the total joint variance of unemployment rates and these three series are accounted for by short-term periodic fluctuations ranging in duration from 2.35 to 2.00 quarters. These results provide clear support for the notion that the earnings of workers in more advantaged socioeconomic structural locations—that is, well educated, in core industries, and associated with union bureaucracies—are protected from the effects of short-term economic fluctuations while earnings of their counterparts are significantly affected by these same short-term fluctuations.

Comparisons of the results for blacks and whites in Table 4 do not support the notion that whites are any more protected from variations in the supply of workers than are blacks. In fact, earnings of whites are more strongly related to all seven of the periodic components with durations given across the top of Table 4.[5] Table 5

[5] Blacks did earn less than whites in this particular survey; for the June 1970 quarter, the mean earnings for whites was $1000.00 greater than the mean for blacks.

indicates the substantial proportions of total joint variance in unemployment rates and earnings are associated with short-term periodic components for both blacks and whites. Possible explanations for these particular findings are not easily discerned, although one possibility is that although blacks may be more often prevented from entering the labor market because of discriminatory hiring practices. Those who do secure employment are no more or less protected from the effects of periodic economic fluctuations than are whites. Other possible explanations for the similarity in the relationship for blacks and whites may be related to the racial composition of the Detroit area.[6] However, results for blacks and whites do lend further support for the general hypothesis that individual earnings patterns are related to unemployment rates via short-term periodic components.

CONCLUSIONS AND IMPLICATIONS FOR FURTHER RESEARCH

This analysis yields substantial support for a relationship between individual earnings and periodic economic fluctuations as measured by unemployment rates. That relationship assumes the form of a delay model, in that changes in employment rates are followed by changes in individual earnings patterns. Thus, macroeconomic-based predictions about the nature of a relationship between structural economic fluctuations and individual earnings have been substantiated to a large degree. However, these results also largely substantiate the notion that individuals who occupy different positions in the social and economic structure of society will be differentially successful in dealing with the effects of periodic economic fluctuations. Indeed, the very existence of the general relationship between individual earnings patterns and unemployment rates is largely the result of socioeconomic structural location. In addition, this analysis substantiates the claim of structural theorists that economic institutional positions are at least as relevant as individual characteristic positions for predicting the nature of structural constraints on individual earnings patterns.[7]

Much remains to be discovered concerning the precise nature and form of these relationships. Clearly, other economic variables measuring short-term economic fluctuations, such as the GNP, productivity, types of goods produced, and the money supply, may be relevant to individual earnings patterns. Other types of social structural positions should be investigated as well. The influences of migration patterns, family size, work and earnings status of other family members, and age have been investigated in more traditional models (Blau and Duncan, 1967), and it could be expected that such variables might condition the relationship between economic fluctuations and individual earnings. In addition, data for female workers, who traditionally have had significantly different career patterns, could indicate substantially stronger relationships between earnings and unemployment rates (see Rosenfeld, 1980). Further differentiation with respect to occupational, industrial, and bureaucratic structural positions is also needed, since the important differences across such groups may lead to variation in the relationship of economic fluctuations and individual earnings. Finally, the results of this study call for improved retrospective and/or longitudinal data sources. For example, the investigation of long-term effects would be greatly facilitated by data collected from older cohorts, aged 50 to 60, so that periods up to forty years could be investigated. Such data, if gathered from larger, more nationally representative samples, could also facilitate comparative analyses across cohorts, historical periods, and industrial compositions.

---

[6] Blacks constituted approximately 19% of the population in the Detroit SMSA in 1970, while for the U.S. as a whole the comparable figure was slightly over 11% (U.S. Census, 1970, Tables 47 and 68).

[7] By individual characteristic position I mean that individual characteristics such as education and race describe certain social structural positions; the characteristics of economic institutions in which individuals exist describe other social structural positions.

REFERENCES

Bartlett, Maurice S.
1966 An Introduction to Stochastic Processes, 2nd Edition. Cambridge: Cambridge University Press.

Beck, Elwood M., Patrick M. Horan, and Charles M. Tolbert II
1978 "Stratification in a dual economy: a sectoral model of earnings determination." American Sociological Review 43:704–20.

Becker, Gary S.
1964 Human Capital. New York: National Bureau of Economic Research.

Bibb, Robert and William H. Form
1977 "The effects of industrial, occupational, and sex stratification on wages in blue-collar markets." Social Forces 55:974–96.

Blau, Peter M. and Otis D. Duncan
1967 The American Occupational Structure. New York: Wiley.

Bloomfield, Peter
1976 Fourier Analysis of Time Series: An Introduction. New York: Wiley.

Bluestone, Barry, William M. Murphy, and Mary Stevenson
1973 Low Wages and the Working Poor. Ann Arbor: Institute of Labor and Industrial Relations, University of Michigan.

Cole, Robert E.
1977 1970 Detroit Area Study: Careers in Detroit. Ann Arbor: Inter-University Consortium for Political and Social Research.

Duncan, Otis D.
1969 "Inheritance of poverty or inheritance of race?" In Daniel P. Moynihan (ed.), On Understanding Poverty. New York: Basic Books.

Durkheim, Emile
1951 Suicide: A Study in Sociology. New York:
[1897] The Free Press.

Fishman, George S.
1969 Spectral Methods in Econometrics. Cambridge, Massachusetts: Harvard University Press.

Granger, Clive W. J. and Hatanaka, M.
1964 Spectral Analysis of Economic Time Series. Princeton, N.J.: Princeton University Press.

Hauser, Robert M.
1980 "Comment on Beck et al., ASR, October 1978 on 'Stratification in a dual economy.'" American Sociological Review 45:702–11.

Henretta, John C. and Richard T. Campbell
1978 "Net worth as an aspect of status." American Journal of Sociology 83:1204–23.
1976 "Status attainment and status maintenance: a study of stratification in old age." American Sociological Review 41:981–92.

Mayer, Thomas F. and William R. Arney
1974 "Spectral analysis and the study of social change." In H. L. Costner (ed.), Sociological Methodology 1973–1974. San Francisco: Jossey-Bass.

Miller, Roger L.
1973 Economics Today. New York: Harper and Row.

Mincer, Jacob
1974 Schooling, Experience and Earnings. New York: National Bureau of Economic Research.

Parnes, Herbert S.
1970 The Pre-Retirement Years: A Longitudinal Study of the Labor Market Experience of Men. Manpower Research Monograph No. 15. 2 vols. Washington, D.C.: U.S. Department of Labor.

Rosenfeld, Rachel A.
1980 "Race and sex-differences in career dynamics." American Sociological Review 45:583–609.

Sewell, William H. and Robert M. Hauser
1972 "Causes and consequences of higher education: models of the status attainment process." American Journal of Agricultural Economics 54:851–61.

Spilerman, Seymour
1977 "Careers, labor market structure, and socioeconomic achievement." American Journal of Sociology 83:551–93.

Tintner, Gerhard
1940 The Variate Difference Method. Bloomington, Indiana: Principia Press.

U.S. Bureau of the Census
1970 Census of the Population. Part 1, Tables 47, 68. Washington, D.C.: Government Printing Office.

U.S. Department of Labor
1955–1961 The Labor Market and Employment Security. Washington, D.C.: Bureau of Employment Security.
1962–1970 Area Trends in Employment and Unemployment. Washington, D.C.: Bureau of Employment Security, Manpower Administration.

DIAGONAL MOBILITY MODELS: A SUBSTANTIVELY MOTIVATED CLASS OF DESIGNS FOR THE ANALYSIS OF MOBILITY EFFECTS*

MICHAEL E. SOBEL
Office of Research and Evaluation, Bureau of Labor Statistics

Designs for the analysis of mobility effects are flawed by the manner in which the effects of origins and destinations are parametrized. Consequently, suggestions that mobility experiences do not affect attitudes and behaviors are inconclusive. A new class of models (diagonal mobility models) for the analysis of mobility effects which are demonstrably grounded in sociological theory is proposed, and estimation of these models is discussed. To illustrate the use of these models, a reanalysis of the relationship between fertility and intergenerational mobility is presented, using data from the 1962 Occupational Changes in a Generation (OCG-1) survey. The effects of origin and destination on fertility do not depend upon particular origin or destination classes; the effect of destination on fertility appears to be stronger than the effect of origin; and there is only weak evidence of mobility effects.

In 1966 Duncan proposed a new strategy for the analysis of mobility effects, arguing that an effect of mobility is to be distinguished from the main effects, arguing that an effect of mobility is to be distinguished from the main effects of the variables that are used to define mobility. This touched off a wave of research that yielded a substantial body of evidence that mobility does not, per se, influence attitudes and behaviors (Laslett, 1971; Knoke, 1973; Jackman, 1972; Kessin, 1971; Boyd, 1971; Jackson and Curtis, 1972; Blau and Duncan, 1967).

Hope (1971, 1975) later argued that Duncan's model (hereafter the square additive or SA model) is seriously flawed by the manner in which the main (nonmobility) effects are parametrized, thus invalidating previous evidence. To correct for the deficiencies of the SA model, Hope (1975) proposed an alternative design, the diamond additive (or DA) model. Though not unquestioned (Halaby and Sobel, 1979; House, 1978), the diamond additive model has yet to be subjected to a thorough analysis. I shall demonstrate that the diamond additive model does not correspond to Hope's own statements about mobility. This disjuncture is serious enough to warrant the conclusion that the DA model, like the SA model, fails to parametrize the main (nonmobility) effects so as to permit the assessment of mobility effects. Thus, a new model that avoids the pitfalls of the SA and DA models must be constructed.

This paper proposes such a model (the diagonal mobility model) for the analysis of mobility effects. Unlike its predecessors, the diagonal mobility model parametrizes the main effects to maintain the correspondence with substantive sociological theory. Thus, mobility effects may be assessed net of a meaningful baseline model. Limitations of the model are discussed in the concluding section.

One caveat: Whereas the SA and DA models have been used in both mobility and status inconsistency research, the

---

* Although the designs proposed are called diagonal mobility models, these should not be confused or equated with Goodman's (1972) diagonal mobility models, which arise in a different substantive context.

Direct all correspondence to Michael E. Sobel, Department of Sociology, Indiana University, Bloomington, Indiana 47405. For detailed comments and suggestions on an earlier version of this paper I wish to thank Neil Fligstein, William Form, Robert M. Hauser, and two anonymous referees. For his assistance in obtaining the solution A.7–A.9 to the system A.4–A.6, I thank Richard J. McDonald. In addition, I am indebted to Donald Treiman, Keith Hope, and Richard Curtis for their comments, and to Charles Halaby for encouraging these interests. The opinions expressed in this paper are the sole responsibility of the author, and do not reflect the policies of the Bureau of Labor Statistics or the views of other BLS staff members.

Table 1. 4 × 4 Mobility Tables for Square and Diamond Additive Mobility Models

1A: Square

	Destination			
Origin	1	2	3	4
1	a_{11}	a_{12}	a_{13}	a_{14}
2	a_{21}	a_{22}	a_{23}	a_{24}
3	a_{31}	a_{32}	a_{33}	a_{34}
4	a_{41}	a_{42}	a_{43}	a_{44}

1B: Diamond

	General Status						
Steps	2	3	4	5	6	7	8
−3				a_{14}			
−2			a_{13}		a_{24}		
−1		a_{12}		a_{23}		a_{34}	
0	a_{11}		a_{22}		a_{33}		a_{44}
1		a_{21}		a_{32}		a_{43}	
2			a_{31}		a_{42}		
3				a_{41}			

diagonal mobility model is a mobility model, and I do not claim that it is necessarily appropriate to the study of status inconsistency.[1]

THE FAILURE OF THE SA AND DA MODELS

For Table 1 the SA model (without mobility effects) parametrizes the observations y_{ijk}; $i=1, \ldots, T$; $j=1, \ldots, T$; $k=1, \ldots, n_{ij}$; where n_{ij} is the number of observations in the ijth cell of the mobility table, as

$$y_{ijk} = B_{1.} + \sum_{q=2}^{T} B_{q.} X_{q.} + \sum_{m=2}^{T} B_{.m} X_{.m} + \varepsilon_{ijk} \qquad (1.1)$$

where $B_{1.} = \mu_{11}$, the population mean in the 11th cell; $X_{q.} = 1$ if $q=i$, 0 otherwise; $X_{.m} = 1$ if $j=m$, 0 otherwise; and ϵ_{ijk} is an observation specific stochastic term with 0 expectation.

To test for mobility effects, Duncan (1966, 1967) suggests a comparison between the baseline model (1.1) and the expanded model which contains row by column interactions. Mobility (status discrepancy) effects are said to exist when the increment to explained variance is statistically significant, and when "these deviations are in some systematic way related to the notion of mobility" (Blau and Duncan, 1967:377). As an alternative to this unwieldy criterion, Hope (1971) identifies three properties of mobility—existence, direction, and magnitude—and demonstrates how to estimate parameters that correspond to these properties.

In more recent research, Hope (1975) argues that the SA model parametrizes the main (nonmobility) effects so as to preclude inferences about the effects of mobility. First, the number of steps moved, a mobility variable, is a linear combination of the variables used to define origins and destinations in the SA model; that is,

$$(i-j) = \sum_{q=2}^{T} (q-1) X_{q.} - \sum_{m=2}^{T} (m-1) X_{.m} .$$

Hence, the SA model "cannot test for the presence of a mobility (status-discrepancy) effect because it incorporates such an effect within its own variance" (Hope, 1975:332). Second, the origin and destination effects cannot be meaningfully equated with the main effect parameters from an analysis of variance, as in the SA model. According to Hope (1971:1025, 1975:336) the characteristic orientation of a class is determined by the stayers: thus the main effect parameters should be determined by the diagonal cells of the mobility table.

As an alternative design for the analysis of mobility effects (or inconsistency effects) Hope (1975) proposes the diamond additive model (Table 1B). The DA model parametrizes the observations y_{ijk} as

[1] This departure from prior treatments (Duncan, 1966:91; Hope, 1975:330) is based upon rejection of the mobility-inconsistency analogy. The rationale for the analogy, namely that mobility may be regarded as temporal inconsistency (Hope, 1975:330) is weak and ignores conceptual and substantive differences between the two phenomena which may prove more important than any superficial similarities. In fact, Hope (1975) fails to maintain the analogy in his verbal formulation of the issues, despite the fact that he imposes the same statistical model on analyses of mobility and inconsistency.

$$y_{ijk} = \delta_{.2} + \sum_{q+m=3}^{2T} \delta_{.(q+m)} Z_{.(q+m)} + \sum_{q-m \neq 0} \delta_{(q-m).} Z_{(q-m).} + v_{ijk} \tag{1.2}$$

where $\delta_{.2} = u_{11}$, $Z_{.(q+m)} = 1$ if $(q+m) = (i+j)$, 0 otherwise; $Z_{(q-m).} = 1$ if $(q-m) = (i-j)$, 0 otherwise; and v_{ijk} is an observation specific stochastic term with 0 expectation. In this model the column parameters $\delta_{.(q+m)}$ represent the effects of the "overall measure of general status" (Hope, 1975:328) while the row parameters $\delta_{(q-m).}$ may be equated with mobility.

In the DA model, the column variables are constructed with respect to two status measures. When the status axes are identical, as in mobility research, each column variable represents a binary contrast between a particular sum of the origin and destination ranks and the other sums. For example, suppose that the status variable is occupational class and that there exist four occupational classes, ranked 1 to 4. To obtain a sum of 4, one combines origin rank 1 with destination rank 3, origin rank 3 with destination rank 1, and origin rank 2 with destination rank 2. In substantive terms, the DA model lumps together in the same column farmers who become managers with managers who become farmers, and depending upon the sum of the levels, perhaps others. Furthermore, for every sum, there is one nonmobile cell, while for every odd sum, there are no nonmobile cells.

It is clear that this parametrization of the columns is not consistent with Hope's contention that class effects should be parametrized with respect to the diagonal cells of the mobility table. If the claim itself is taken seriously, the DA model is flawed by the same conceptual failure as the SA model: it fails to establish a correspondence between sociological substance and statistical method.

There is, however, a correspondence between the construction of the column variables and a concept I shall call "permanent status," analogous to the economic concept of permanent income. This point (in a different context) has also been recognized by Hope, who, using the term "lifetime class," apparently rejects his earlier position (Hope, 1981:30). Thus, it might be argued that the mobility effects in the DA model represent the effect of deviations from some permanent status on the criterion in question. In this case, the parametrization is only as meaningful as the concept itself, so it is of value to take up this issue; unfortunately, Hope himself does not do so.

In economic work, the concept of permanent income is useful in explaining a certain class of behaviors under a restrictive set of conditions. First, it must be assumed that the individual (household) knows his (its) income trajectory, and is able, in some fashion, to distinguish between transitory and permanent components. Next, it must be assumed that the behavior in question is affected by the level of permanent income, rather than by transitory income, or by the sum of permanent and transitory components. In other words, the individual is able to calculate his permanent income, does so, and adjusts the particular criterion in accordance with it.

There are variety of problems with the analogy between permanent status (lifetime class) and permanent income. Most importantly, there is no compelling theoretical or empirical reason to believe that even the rare individual who is able to predict his occupational trajectory calculates a permanent status which he uses to equilibrate a whole host of behaviors and attitudes over a lifetime. In fact, such a conceptualization would appear to be at odds with the bulk of modern sociological thought.

THE DIAGONAL MOBILITY MODELS[2]

Mobility theory (Blau, 1956; Duncan, 1966; Blau and Duncan, 1967) suggests that an individual's current attitudes, values, and behaviors are influenced by both prior and current status. Mobility

[2] The exposition of this section is simplified, without loss of generality, by presenting the various baseline models, as opposed to the analogues which incorporate mobility effects.

effects are effects over and above this partial determination, i.e., mobility effects are those systematic influences which are left after the process of acculturation (socialization to reference norms) has been modeled. Thus, a meaningful mobility model must parametrize the acculturation process in a sociologically sensible fashion. Substantively, this requires the choice of a referent; acculturation is a social process whereby individuals adopt the relevant behaviors (values, attitudes) which typify the reference aggregate. Thus, there must be some values of the dependent variable which typify the appropriate referents.

Consider the additive baseline model (simple diagonal model without mobility effects):[3]

$$y_{ijk} = p\mu_{ii} + r\mu_{jj} + \epsilon_{ijk};$$

$$i=1,\ldots,T;\ j=1,\ldots,T;$$

$$k=1,\ldots,n_{ij}; \tag{2.1}$$

$$p + r = 1;\ \text{i.e., } r = 1 - p; \tag{2.2}$$

$$p \text{ lies in } [0, 1]; \tag{2.3}$$

where ϵ_{ijk} is a stochastic term with 0 expectation, and μ_{ii} and μ_{jj} are the population means in the iith and jjth cells of the mobility table.

In this model, the referent values are chosen with respect to the diagonal of the mobility table, as suggested by Hope (1971, 1975). For observations on the diagonal, the referent value is μ_{ii} and Equations 2.1 and 2.2 reduce to

$$y_{iik} = \mu_{ii} + \epsilon_{iik};\ i=1,\ \ldots,\ T;$$

$$k=1,\ \ldots,\ n_{ii}.$$

For observations off the diagonal, two referent values are given. The first is given by the diagonal mean of the origin category, the second by the diagonal mean of the destination category. If an individual starts in origin category i, it is assumed that the appropriate prior reference aggregate consists of stayers in the ith category, with mean μ_{ii}; if that individual moves to the jth destination category, his other referent is comprised by the stayers in the jth category, with means μ_{jj}.

But both referents figure into the final determination, in accord with the notion that the outcome is a function of both prior and current status. Equation 2.1 takes this into account by assigning a parameter to origin status and a parameter to destination status, subject to the equality constraint that the parameters sum to one (Equation 2.2). Imposition of the constraint that p lie in the closed unit interval (Equation 2.3) insures that the parameters p and $r = 1 - p$ may be interpreted as origin and destination weights, respectively. Thus, the parameter p indicates the salience of origin status, relative to that of destination status, to the determination of the phenomenon under investigation.

Unlike the DA model, and like the SA model, the conceptual distinction between origins and destinations is preserved in the diagonal model. But in contrast to the SA model, the maintenance of this distinction does not preclude any inferences about the effects of mobility. The model given by Equations 2.1–2.3 does not already incorporate mobility effects due to existence, direction, steps, or higher order polynomials in mobility. Hence, mobility variables and parameters may be added to the model to unambiguously test against the null hypothesis of acculturation.

The simple diagonal model thus satisfies the two minimal criteria. In addition, like the SA model, there is no assumption about the order of the categories in Equations 2.1–2.3; that is, unlike the DA model, the diagonal model is applicable even when the ordering of the categories is unknown. Finally, the diagonal model does not require that $\mu_{ij} = \mu_{ii}$ when $i \neq j$; but if p happens to be .5, then $\mu_{ij} = \mu_{ji}$ if no mobility effects are present, and the

[3] In general, theoretical considerations do not suggest a clear preference for the additive model, as opposed to the multiplicative model:

$$Y_{ijk} = \left(\sum_{q=1}^{T} \mu_{qq}X_{q.}\right)^{p}\left(\sum_{m=1}^{T} \mu_{mm}X_{.m}\right)^{1-p} \exp\{\epsilon_{ijk}\},$$

where the notation is as previously described, and $0 \leq p \leq 1$, as before. If the dependent variable is a positive random variable, taking logarithms yields:

$$\log y_{ijk} = p \log \sum_{q=1}^{T} \mu_{qq}X_{q.} + (1-p) \log \sum_{q=1}^{T} \mu_{mm}X_{.m} + \epsilon_{ijk}.$$

Thus, the multiplicative form is additive in the logarithms, and does not require independent exposition.

resulting model yields the same expectations as Hope's (1971) halfway model. In this vein, it should also be noted that the simple diagonal model subsumes Hope's (1975:335) model d.[4]

However, the new baseline model is open to the objection that it is overly restrictive to require that the same proportionality constant characterize all origin and destination classes. The objection may be addressed by comparing Equations 2.1–2.3 with the additive diagonal mobility model (DM-1) given by:

$$y_{ijk} = p_{i.}\,\mu_{ii} + r_{i.}\,\mu_{jj} + \varepsilon_{ijk};$$

$$i=1,\ldots,T;\; j=1,\ldots,T;$$

$$k=1,\ldots,n_{ij}; \tag{2.4}$$

$$p_{i.} + r_{i.} = 1;\; \text{i.e.,}\; r_{i.} = 1 - p_{i.};$$

$$i=1,\ \ldots,\ T; \tag{2.5}$$

$$p_{i.} \text{ lies in } [0,\ 1];\; i=1,\ldots,\ T \tag{2.6}$$

where $p_{i.}$ is an origin-specific proportionality constant and the other notation is as previously described. In the next section, I prove that the simple model is nested under the DM-1 model; this fact allows the comparison between Equations 2.1–2.3 and 2.4–2.6.

Unlike the simple diagonal mobility models, the effect of destination in the diagonal mobility models, because it is origin specific, varies within the particular destination. Thus, an individual who moves from class i to class j and an individual who moves from class i + 1 to class j are characterized by different origin and destination effects, unless $p_{i.} = p_{(i+1).}$. While this property of the model might seem unusual, it merely restates the sensible equality constraints which require that, for any individual, the proportions sum to one.

While the equality constraints are reasonable, and necessary for a meaningful substantive interpretation of the model, they are also responsible for a rather curious asymmetry. The diagonal mobility models condition on origins and measure the effect of destination as a deviation from the origin effect. Conditioning on destinations yields the model DM-2:

$$y_{ijk} = p_{.j}\mu_{jj} + r_{.j}\mu_{ii} + \varepsilon_{ijk};$$

$$i=1,\ldots,T;\; j=1,\ldots,T;$$

$$k=1,\ldots,\ n_{ij}; \tag{2.7}$$

$$p_{.j} + r_{.j} = 1;\; \text{i.e.,}\; r_{.j} = 1 - p_{.j};$$

$$j=1,\ldots,T; \tag{2.8}$$

$$p_{.j} \text{ lies in } [0,\ 1];\; j=1,\ldots,T; \tag{2.9}$$

where $p_{.j}$, j=1, . . ., T is a destination specific proportionality constant.

Now suppose that estimation of either the DM-1 or DM-2 models from sample data is not problematic. Then the asymmetry arises because it is not necessary that $\hat{p}_{m.} = (1-\hat{p}_{.m})$ and thus it is possible to reach substantive conclusions which depend upon whether origin class or destination class is conditioned on. For example, if m = 1, then $\hat{p}_{1.}$ need not equal $(1 - \hat{p}_{.1})$, even though both estimates measure the salience of the first origin category.[5]

ESTIMATION OF THE DIAGONAL MOBILITY MODELS

While the diagonal mobility models are mathematically simple, the estimation problem is not trivial. (This is a highly

[4] I thank Keith Hope for this observation.

[5] In any given context, it is not immediately clear that either the DM-1 or DM-2 model is preferable to the analogous simple diagonal mobility model. This is an empirical issue. If the simple model suffices, then the asymmetry in the analogous diagonal mobility model is not problematic. Furthermore, if substantive preferences for conditioning on origin status or destination status can be established, the asymmetry need cause little concern. In fact, only when (1) one of the simple models is inadequate and (2) the researcher is unable to establish substantive preferences for conditioning on origin or destination and (3) when the inferences derived from the DM-1 and DM-2 models about the main and/or mobility effects conflict is the asymmetry a cause for concern. I suspect that in the typical case this particular conjunction is unlikely. But when the first two elements of the conjunction occur, it is recommended that researchers estimate both DM models and draw definite conclusions only when the analogous features of the two models are in accord.

technical section: readers who are not interested in the estimation procedures can proceed to the next section.) Below I focus on the simple diagonal mobility model, with mobility effects; the results generalize easily to the DM-1 and DM-2 models, and to their multiplicative analogues, and the discussion subsumes the cases where no mobility effects are present.[6]

The simple diagonal model with mobility effects is given by

$$y_{ijk} = p\mu_{ii} + r\mu_{jj} + \sum_{w=1}^{W} \gamma_w x_{ijw} + \varepsilon_{ijk};$$

$$i=1,\ldots,T;\ j=1,\ldots,T;$$

$$k=1,\ldots,n_{ij}; \qquad (3.1)$$

$$p + r = 1;\ \text{i.e., } r = 1 - p; \qquad (3.2)$$

$$p \text{ lies in } [0, 1] \qquad (3.3)$$

where the γ_w, $w=1, \ldots, W$, are mobility parameters, the x_{ijw} are mobility variables (for example, mover-stayer contrasts, upward vs. downward contrasts, the number of steps moved through the mobility hierarchy, higher order polynomials in steps), and the other notation is as previously described.

To develop the results of interest, Equation 3.1 must be expressed in an alternative form. To this end, as before, let $X_{q.} = 1$ if $q = i$, 0 otherwise; and let $X_{.m} = 1$ if $j = m$, 0 otherwise.

Define

$$\mu_1 = \sum_{q=1}^{T} \mu_{qq} X_{q.} \quad \text{for all } i, j, k;$$

and let

$$\mu_2 = \sum_{m=1}^{T} \mu_{mm} X_{.m} \quad \text{for all } i, j, k.$$

Equation 3.1 then takes the form:

$$y_{ijk} = \sum_{q=1}^{T} p\mu_{qq} X_{q.} + \sum_{m=1}^{T} (1-p)\mu_{mm} X_{.m}$$

$$+ \sum_{w=1}^{W} \gamma_w x_{ijw} + \varepsilon_{ijk}; \qquad (3.4)$$

Equation 3.4 now incorporates the equality constraint of 3.2. Next, let $\underline{\mu}_1$ and $\underline{\mu}_2$ be the column vectors of μ_1 and μ_2, let $\underline{y}$ be the vector of y_{ijk}, let $\underline{\epsilon}$ be the vector of ϵ_{ijk}, let X be a matrix of mobility variables, and let $\underline{\gamma}$ be a vector of mobility parameters. Equation 3.4 may now be expressed in matrix form as

$$\underline{y} = (\underline{\mu}_1, \underline{\mu}_2, X) \begin{pmatrix} p \\ 1-p \\ \underline{\gamma} \end{pmatrix} + \underline{\varepsilon} \qquad (3.5)$$

which is equivalent to

$$(\underline{y} - \underline{\mu}_2) = (\underline{\mu}_1 - \underline{\mu}_2, X) \begin{pmatrix} p \\ \underline{\gamma} \end{pmatrix} + \underline{\varepsilon}. \qquad (3.6)$$

To the model comprised by Equations 3.6 and 3.3 we now add two final assumptions:

$$(\underline{\varepsilon} \mid \underline{\mu}_1, \underline{\mu}_2, X) \sim N(\underline{0}, \sigma^2 I) \qquad (3.7)$$

$$(\underline{\mu}_1, \underline{\mu}_2, X) \text{ has full column rank.}^7 \qquad (3.8)$$

If the vectors $\underline{\mu}_1$ and $\underline{\mu}_2$ were observed, ordinary least squares (OLS) estimation of the model given by Equations 3.3 and 3.6–3.8 yields estimators which are best linear unbiased because the restriction 3.3 does not reduce the dimension of the parameter space (Rothenberg, 1973:55). Furthermore, if the OLS estimate of p satisfies 3.3, OLS is equivalent to maximum likelihood.

However, $\underline{\mu}_1$ and $\underline{\mu}_2$ are unobserved, and thus it is not entirely clear how to

[6] More precisely, the generalization to the multiplicative models holds when the logarithmically additive form is estimated, under the assumption that the disturbances from this equation, given the explanatory variables, are distributed independently according to a $N(0, \sigma^2)$ distribution.

[7] In the SA model, the matrix of mobility variables cannot simultaneously include the number of steps moved and satisfy condition 3.8. In the DA model, other restrictions on the form of X are imposed (see Halaby and Sobel, 1979). But the diagonal model satisfies (3.8) under the restrictions that X itself has full rank and that μ_1 is not a linear combination of μ_2 and X. These restrictions are much weaker than those needed to identify either the SA or DA models. The diagonal model is therefore identifiable under very general conditions.

estimate the model. The naive way is to estimate μ_1 and μ_2 by

$$\sum_{q=1}^{T} \bar{y}_{qq} X_{q.} \quad \text{and} \quad \sum_{m=1}^{T} \bar{y}_{mm} X_{.m} ,$$

respectively, where $\bar{y}_{qq}$ and $\bar{y}_{mm}$ are the sample means of the *qq*th and *mm*th cells of the mobility table, and to use these as proxies for μ_1 and μ_2 in the regression problem, assuming an interior solution (i.e., assuming that the restriction 3.3 is satisfied). It is easy to show that this procedure yields biased but consistent parameter estimates. However, it is clear that the variance-covariance matrix of the disturbances is no longer diagonal, and ordinary least squares ignores this information. Furthermore, employment of this procedure requires that each diagonal cell contain at least one observation. Thus, this procedure is unattractive.

Because the form of the error structure is known (see Equation 3.7, the likelihood function under the simple model, assuming random sampling, may be written as

$$\prod_{i,j,k} ((2\pi\sigma^2)^{-1/2} \exp\{-(2\sigma^2)^{-1} \times (y_{ijk} - \sum_{q=1}^{T} p\mu_{qq} X_{q.} - \sum_{m=1}^{T} (1-p) \mu_{mm} X_{.m} - \sum_{w=1}^{W} \gamma_w X_{ijw})^2\}).$$

It is simple to show that maximization of the log likelihood function with respect to p; μ_{ii}, $i = 1, \ldots, T$; γ_w, $w = 1, \ldots, W$; σ^2 is equivalent to choosing these parameters so as to minimize the function

$$\sum_{i,j,k} (y_{ijk} - \sum_{q=1}^{T} p\mu_{qq} X_{q.} - \sum_{m=1}^{T} (1-p) \mu_{mm} X_{.m} - \sum_{w=1}^{W} \gamma_w x_{ijw})^2 .$$

That is, maximum likelihood estimation and nonlinear least squares are equivalent. (For further discussion of the estimation procedures, see Appendix.) Thus, simultaneous estimation of the parameters p, μ_{ii}, $i=1, \ldots, T$, and γ_w, $w=1, \ldots, W$, by nonlinear least-squares yield estimates which, under general regularity conditions, are consistent, asymptotically normal, and asymptotically efficient (Theil, 1971:395–7; Goldfeld and Quandt, 1972:57–74; Malinvaud, 1970:338–41; Rao, 1973:353–66). It does not, as I show in the appendix, yield estimates of the μ_{ii}, $i=1, \ldots, T$, that are based solely on the diagonal cell means, although this is not a disadvantage of the procedure.[8]

The simple diagonal model, with or without mobility effects, is nested under the analogous DM-1 or DM-2 model. To see this, we differentiate the n elements of $\underline{y}$ with respect to p, yielding the vector $\underline{\mu}_1 - \underline{\mu}_2$. Similarly, differentiation of $\underline{y}$ with respect to $p_{q.}$, $q=1, \ldots, T$ under the analogous DM-1 model yields, for the *ijk*th observation, the row vector:

$$(\mu_{11} X_{1.} - (\sum_{m=1}^{T} \mu_{mm} X_{.m}) X_{1.}, \ldots, \mu_{TT} X_{T.} - (\sum_{m=1}^{T} \mu_{mm} X_{.m}) X_{T.}).$$

Summing the elements and arranging these sums into a column vector yields $\underline{\mu}_1 - \underline{\mu}_2$, the first column of the matrix

$$X^* = (\partial \underline{y}/\partial p,\ \partial \underline{y}/\partial \mu_{11},\ \ldots,\ \partial \underline{y}/\partial \mu_{TT},\ \ldots, \partial \underline{y}/\partial \gamma_W).$$

[8] Friendly critics have pointed out that it seems contradictory to criticize Hope for mixing diagonal and off-diagonal cells in his construction of the column variables when my estimates of the μ_{ii} are not a function of only the diagonal means. However, this contradiction is more apparent then real: once the distinction between the population and the sample is kept in mind, there is no contradiction. In the population, Hope's model mixes diagonal and off-diagonal cells, whereas the diagonal model parametrizes the main effects with respect to the diagonal population means. In the sample, the diagonal population means must be estimated. The procedure utilized is maximum likelihood, and it turns out, under the model, that the maximum likelihood estimators of the diagonal population means are not solely a function of the diagonal sample means. Problems arise if the diagonal sample means are used to estimate the diagonal population means and ordinary least squares is then used to estimate the other parameters. By using maximum likelihood, these problems are alleviated and estimates for the diagonal population means that are "best asymptotically normal" under the model are obtained. Thus, to argue that estimates of the diagonal population means should be based on the diagonal sample means is to argue against the principle of maximum likelihood estimation.

Table 2. Mean Number of Children Ever Born, by Husband's Father's Occupation and Husband's 1962 Occupation, for Wives 42 to 61 Years Old in March, 1962, Living with Husband, in OCG Sample

Husband's Father's Occupation	Husband's Occupation in 1962: White Collar		Manual		
	Higher	Lower	Higher	Lower	Farm
Higher white collar	2.024	1.752	2.031	2.308	2.850
	(538)	(137)	(129)	(130)	(20)
Lower white collar	2.005	1.631	2.389	2.225	2.700
	(190)	(65)	(54)	(71)	(10)
Higher manual	2.113	2.020	2.407	2.349	1.900
	(318)	(153)	(307)	(269)	(20)
Lower manual	2.180	1.908	2.423	2.423	3.000
	(322)	(163)	(345)	(513)	(34)
Farm	2.252	2.142	2.760	2.850	3.194
	(389)	(148)	(462)	(687)	(484)

NOTE: Numbers in parentheses are the number of observations within cells; n = 5958.

But X* is simply the nonlinear analogue to the design matrix in linear regression (Gallant, 1975), and so the result shows that the vectors in the design matrix for the simple model span a subspace of the space spanned by the vectors in the analogous DM-1 model. Similarly, the simple model is nested under the analogous DM-2 model.

ARE THERE MOBILITY EFFECTS IN THE OCG-I DATA?

As an example of the use of the diagonal mobility models, the relationship between intergenerational mobility (father's occupation to son's occupation in March 1962) and fertility (number of children ever born per wife) is reexamined with the OCG-I data (Blau and Duncan, 1967). The same population of interest (wives 42 to 61 years of age in March 1962, living with OCG husbands) was selected for study, and the occupations were aggregated by the same classification scheme (Blau and Duncan, 1967:371–2) into a 5 × 5 mobility table, with one exception: Blau and Duncan include a "not stated" category in their analysis, and this category is omitted from the present analysis. Category 1 contains higher white-collar workers, category 2 contains lower white-collar workers, 3 contains higher manual workers, 4 contains lower manual workers, and category 5 contains farmers.

Preliminary analysis indicates (a) that it is acceptable to focus on the additive version of the models; (b) that a more acceptable fit is obtained in analyses which use the variable STEPS = (i − j), rather than its absolute value; and (c) that the contribution of higher order mobility effects (e.g., $STEPS^2$) is negligible. Thus, the mobility variables considered are (a) MOB, a mover-stayer contrast which takes on the value of 1 for movers, 0 for stayers; (b) DIR, a variable equal to 0 for stayers, 1 for movers up, −1 for movers down, and (c) STEPS, a variable that equals i − j, the difference between origin and destination ranks.[9]

Table 2 presents the n_{ij} and $\bar{y}_{ij}$ for the 5 by 5 table. Interested readers may wish to compare these $\bar{y}_{ij}$ with those predicted under the models considered here. For example, to compare the $\bar{y}_{ij}$ with the fitted values under the simple diagonal model without mobility effects one uses the estimates $\hat{p}$ and $\hat{\mu}_{ii}$, i=1, . . ., T, to compute $y_{ijk} = \hat{p}\,\hat{\mu}_{ii} + (1 - \hat{p})\,\hat{\mu}_{jj}$; this yields the predicted mean in the *ij*th cell.

Table 3 presents the estimates for three versions of the simple diagonal model, and

[9] Of course, any model that contains the variable STEPS may be considered conceptually equal-interval. Similarly, any model that contains DIR, but not STEPS, assumes that the ranking of the categories is known. But if only MOB is included, no order assumptions are introduced at all.

Table 3. Coefficients and Asymptotic Standard Errors for Parameters of the Simple Diagonal Mobility Models

	Simple, no mobility effects	Simple	Simple final
$\hat{p}$	.3551	.3618	.3597
	(.0515)	(.1022)	(.1000)
$\hat{\mu}_{11}$	1.9697	1.9984	1.9948
	(.0628)	(.0756)	(.0690)
$\hat{\mu}_{22}$	1.6903	1.6867	1.6803
	(.1111)	(.1190)	(.1134)
$\hat{\mu}_{33}$	2.3993	2.3793	2.3741
	(.0719)	(.0859)	(.0730)
$\hat{\mu}_{44}$	2.4849	2.4587	2.4542
	(.0621)	(.0755)	(.0648)
$\hat{\mu}_{55}$	3.2006	3.1952	3.1931
	(.0812)	(.0834)	(.0817)
MOB—$\hat{\gamma}_1$		−.0074	
		(.0637)	
DIR—$\hat{\gamma}_2$		.1626	.1618
		(.0705)	(.0702)
STEPS—$\hat{\gamma}_3$		−.0678	−.0678
		(.0440)	(.0440)
$\hat{\sigma}^2$	3.9626	3.9591	3.9591

NOTE: MOB = 0 if i = j; 1 otherwise.
DIR = 1 if i > j, 0 if i = j, −1 if i < j.
STEPS = i − j.
n = 5958; standard errors (under simple random sampling) enclosed in parentheses.

Table 4 presents estimates for the DM-1 and DM-2 models, with and without mobility effects. One cautionary note: to compare the parameters of the DM-1 and DM-2 models in Table 4, remember that $p_{i.}$ measures the salience of origins in the DM-1 model and $p_{.j}$ measures the salience of destinations in the DM-2 model (see Equations 2.4 and 2.7). Thus, $\hat{p}_{1.}$ should be compared with $1 - \hat{p}_{.1}$, etc.

Tables 3 and 4 indicate that inclusion of mobility effects causes the standard errors of the other estimates to increase and the estimates often increase in magnitude as well. The estimates $\hat{\gamma}_1$ and $\hat{\gamma}_3$ are not significant at the .05 level (based on the N(0,1) distribution) in any of the models, but $\hat{\gamma}_2$ is significant in the simple diagonal and DM-1 models, although it is not significant in the DM-2 model. In the DM-1 and DM-2 models the parameter estimates are often too discrepant across models to permit strong conclusions about the exact effects of starting in a particular origin class or winding up in a particular destination class. However, no version of

Table 4. Coefficients and Asymptotic Standard Errors for Parameters of the Additive DM-1 and DM-2 Models

	DM-1 no mobility effects	DM-1		DM-2 no mobility effects	DM-2
$\hat{p}_{1.}$	.4403	.5971	$\hat{p}_{.1}$	.8017	.8248
	(.1978)	(.2442)		(.0962)	(.1939)
$\hat{p}_{2.}$	.1438	.1439	$\hat{p}_{.2}$	.7431	.7733
	(.1722)	(.1749)		(.1477)	(.1766)
$\hat{p}_{3.}$	.6120	.5587	$\hat{p}_{.3}$	.5324	.6054
	(.1708)	(.2221)		(.1217)	(.1573)
$\hat{p}_{4.}$	.3793	.4774	$\hat{p}_{.4}$	.4885	.5780
	(.1456)	(.2301)		(.1122)	(.1428)
$\hat{p}_{5.}$	.3606	.4420	$\hat{p}_{.5}$	.4755	.5049
	(.0616)	(.1230)		(.2263)	(.2412)
$\hat{\mu}_{11}$	1.9596	2.0103	$\hat{\mu}_{11}$	2.0390	2.0128
	(.0655)	(.0746)		(.0646)	(.0759)
$\hat{\mu}_{22}$	1.5764	1.5571	$\hat{\mu}_{22}$	1.7772	1.7517
	(.1466)	(.1875)		(.1495)	(.1637)
$\hat{\mu}_{33}$	2.3586	2.3839	$\hat{\mu}_{33}$	2.3658	2.3527
	(.0686)	(.0894)		(.0843)	(.0945)
$\hat{\mu}_{44}$	2.4923	2.4563	$\hat{\mu}_{44}$	2.4509	2.4435
	(.0673)	(.0778)		(.0720)	(.0754)
$\hat{\mu}_{55}$	3.2155	3.1990	$\hat{\mu}_{55}$	3.2010	3.2035
	(.0870)	(.0877)		(.0872)	(.0890)
MOB—$\hat{\gamma}_1$		−.0272	$\hat{\gamma}_1$		.0172
		(.0728)			(.0682)
DIR—$\hat{\gamma}_2$		.1585	$\hat{\gamma}_2$		.0632
		(.0756)			(.1043)
STEPS—$\hat{\gamma}_3$		−.0872	$\hat{\gamma}_3$		−.0097
		(.0453)			(.0663)
$\hat{\sigma}^2$	3.9606	3.9575	$\hat{\sigma}^2$	3.9580	3.9574

NOTE: n = 5958; standard errors (under simple random sampling) enclosed in parentheses.

either the DM-1 or DM-2 model is clearly superior to its simple diagonal analogue.

To investigate this more formally, we compute the likelihood ratios $\lambda=(\hat{\sigma}_F/\hat{\sigma}_N)$ (Goldfeld and Quandt, 1972:74), where $\hat{\sigma}_N$ is the maximum likelihood estimator of the square root of the error variance in the nested model, $\hat{\sigma}_F$ is the maximum likelihood estimate of the corresponding quantity in the unconstrained or full version of the model. We then use the fact that $-2(\log\lambda)$ has an asymptotic $\chi^2(r)$ distribution, where r is the number of additional independent parameters, in the unconstrained model, to make statistical comparisons among the appropriate models.

In no instance is any DM-1 or DM-2 model statistically superior to its simple analogue at the .1 level of significance.[10] I conclude that there appear to be no differential effects of starting (or ending) in any particular class. Thus it suffices to focus attention on the simple diagonal models. Under all three versions, the cell means are fitted quite well by the model. For example, in the simple model without mobility effects, over 99% of the between-cell sums of squares (not deviated about the mean) is accounted for, and over 92% of the deviated between-cell sums of squares is accounted for.[11] Furthermore, in all three versions the estimates of the origin-effect parameters are close to one another, although the hypothesis that p is greater than or equal to .5 cannot be rejected at the .05 level in the simple models which include mobility effects. On balance it appears that the effect of destination on fertility is more important than the effect of origin; $\hat{p}$ is near .36 in all the simple models, hence $1 - \hat{p}$ is near .64, or the effect of destination is (descriptively) nearly twice the effect of origin.

The diagonal model with all three mobility effects (panel 2 of Table 3) is not statistically superior to the simple diagonal model without mobility effects. This suggests that there are no mobility effects in these data and that the simple model without mobility effects is an adequate summary of the data. However, in the simple model with mobility effects the coefficient $\hat{\gamma}_2$ is statistically significant, indicating that upward mobility, in and of itself, appears to increase, rather than decrease, fertility. This effect appears to be partially counterbalanced by the effect of STEPS; although the coefficient $\hat{\gamma}_3$ is not significant at .05, the ratio of the coefficient to its standard error exceeds 1.5 in magnitude, and thereby lends some credence to the notion that each step up (down) the mobility ladder decreases (increases) fertility by approximately 40% of the direction effect. In the face of this almost significant effect, it becomes difficult to compare the experiences of the nonmobile with those of the mobile.

Table 5. Mobility Contrasts for Comparisons with and across the Diagonal of the Mobility Table

Contrast (H_0)	Estimate	Asymptotic standard deviation
$\gamma_2 + \gamma_3 = 0$	.0940	.0525
$\gamma_2 + 2\gamma_3 = 0$	.0262	.0668
$\gamma_2 + 3\gamma_3 = 0$	−.0416	.1002
$\gamma_2 + 4\gamma_3 = 0$	−.1094	.1396
$2\gamma_2 + 2\gamma_3 = 0$	.1880	.1050
$2\gamma_2 + 3\gamma_3 = 0$	.1202	.1117
$2\gamma_2 + 4\gamma_3 = 0$	.0524	.1335
$2\gamma_2 + 5\gamma_3 = 0$	−.0154	.1644
$2\gamma_2 + 6\gamma_3 = 0$	−.0832	.2003
$2\gamma_2 + 7\gamma_3 = 0$	−.1510	.2839
$2\gamma_2 + 8\gamma_3 = 0$	−.2188	.2791

NOTE: Computations, etc. under simple random sampling.

In order to facilitate these comparisons, the simple model was reestimated after omitting the variable MOB, because the ratio of its coefficient to its standard error is well below unity in magnitude. (The resulting model appears in the third panel of Table 3.) Despite the fact that the STEPS coefficient is not significant, the STEPS variable is retained in this, the final model, on the conventional grounds that its coefficient value is greater in absolute value than its standard error.

In this model, the effect of origin is

[10] Nor are the DM-1 and DM-2 models with mobility effects superior to the DM-1 and DM-2 models without mobility effects. Researchers who wish to pursue the comparisons in the text can use the error variances in Tables 3 and 4 to construct the likelihood ratio tests.

[11] I thank Robert M. Hauser for suggesting these calculations.

again descriptively less than .5, although the null hypothesis that it is greater than or equal to .5 cannot be rejected at the .05 level. Since the coefficient for STEPS is not significant, all contrasts among those who moved in the same direction are insignificant. To compare the diagonal cells with the off-diagonal cells and to compare the off-diagonal cells across the diagonal of the mobility table, the covariance between the estimates of the two mobility parameters is needed; its value is .0021.

Our primary interest is to compare the mobile with the nonmobile. To make a specific comparison, it suffices to test

$$H_0: \ \gamma_2 \text{DIR} + \gamma_3 \text{STEPS} = 0 \text{ against}$$

$$H_1: \ \gamma_2 \text{DIR} + \gamma_3 \text{STEPS} \neq 0.$$

Since the test is symmetric across the diagonal it suffices to consider only those who moved up (see the first four entries in Table 5). Clearly, none of the comparisons are significant at the .05 level; however, the effect for one-step movers has a p value of .07, based on the N(0,1) distribution. Thus, there is weak evidence that one-step movers exhibit mobility effects; movement up increases fertility by .094 children, movement down decreases fertility by .094 children.

To compare cells across the diagonal of the mobility table, it is sufficient to test

$$H_0: \ 2\gamma_2 \text{DIR} + \gamma_3 |\Delta \text{STEPS}| = 0 \text{ against}$$

$$H_1: \ 2\gamma_2 \text{DIR} + \gamma_3 |\Delta \text{STEPS}| \neq 0,$$

where ΔSTEPS is the difference in STEPS between the upwardly and downwardly mobile. The remaining entries in Table 5 present the estimated differences and standard deviations for these comparisons. With the exception of the contrast between those who moved up one step and those who moved down one step (this contrast is equivalent to the contrast between one-step movers and stayers), there are no differences of any import.

Taken as a whole, the analysis itself seems to bear out the conclusions reached by Blau and Duncan (1967:397): "By and large the fertility of mobile couples, which is intermediate between that prevailing in their origin and that prevailing in their destination stratum, can be explained by the additive influence of these two social strata." It does not, however, bear out Blau and Duncan's (1967:397) assertion that long-distance mobility, in and of itself, depresses fertility.

Readers should not conclude that, because the new analysis confirms Blau and Duncan's major conclusions, the diagonal model is only a trivial improvement on the SA model. The two models are conceptually different, and the SA model cannot, in and of itself, yield credible inferences about the effects of mobility because the model itself is flawed. The diagonal model, however, is conceptually suited to the problem at hand. Therefore it is always to be preferred over the SA (or DA) model. That the SA and diagonal models yield similar conclusions in this instance is an isolated empirical fact.

In addition, the diagonal models allow substantive inferences which cannot be derived within the square (or diamond) additive framework. The diagonal models suggest descriptively that origins impact on fertility less than destinations. They also suggest that the relative salience of origins and destinations is invariant across origin and destination classes. By using this information, an extremely parsimonious parameterization of the mobility table is obtained.

SUMMARY AND CONCLUSIONS

Previous designs for the analysis of mobility effects are flawed both methodologically and conceptually. Furthermore, in these designs, it is virtually impossible to interpret the main-effect parameters in a meaningful, substantive fashion.

In this paper a class of models (the diagonal mobility models) that does not share the defects of previous designs is proposed, discussed, and estimated. In contrast with prior designs, the diagonal mobility models allow for an unambiguous substantive interpretation of both main effects and mobility effects within a highly flexible framework. Thus, a broader set of inferences may be derived from this class of models.

Future researchers may find it profitable to extend analyses based on the diagonal mobility models in two fashions. First, there is no inherent reason why comparisons between the simple and more complicated versions of the model need to be undertaken solely by means of global tests. If the difference between specific main-effect parameters is of a priori theoretical interest, researchers may wish to make specific contrasts instead of relying upon a global criterion. Second, mobility researchers rarely consider the effects of variables other than origins, destinations, and mobility. Future researchers will surely wish to consider the introduction of other variables into their models before reaching conclusions about the effects of mobility (Kessler and Cleary, 1980).

But most importantly, future researchers should note that the diagonal model is not without fault, and that further theoretical and methodological developments are necessary. The diagonal models assume that the relevant reference aggregates are the diagonal stayers. But if, for example, the proportion of stayers in a destination class is small, this may decrease the visibility of stayers and diminish their ability to define the sociocultural norms that typify the destination class.[12] This suggests that the diagonal models are overly simple minded because they do not take into account the patterns of mobility in the mobility table. To deal with these and other issues, researchers will have to treat mobility and its concomittants simultaneously, and this will require new substantive and methodological breakthroughs. The diagonal models are a small step in the right direction.

APPENDIX

The purpose of this appendix is (a) to show that in general, the maximum likelihood estimators of μ_{ii}, $i=1, \ldots, T$, are not based solely on the sample means of the diagonal cells, and (b) to discuss computational details of the estimation procedure used in this paper. To see the first point consider the simple diagonal model (without mobility effects) for a 2×2 table. Assuming random sampling, the unrestricted likelihood function may be written as

$$\prod_{ijk}(2\pi\sigma^2)^{-n/2} \times$$

$$\exp\{-(2\sigma^2)^{-1}(y_{ijk}-u_{ijk})^2\}, \qquad (A.1)$$

where n is the sample size, and $\mu_{ijk} = p\mu_{ii} + (1-p)\mu_{jj}$. Thus the log likelihood is given by

$$(-n/2)\log 2\pi\sigma^2 -$$

$$(2\sigma^2)^{-1}\sum_{ijk}(y_{ijk}-\mu_{ijk})^2 . \qquad (A.2)$$

For the special case under consideration (A.2) reduces to

$$(-n/2)\log(2\pi\sigma^2)-(2\sigma^2)^{-1} \times$$

$$[\sum_{k=1}^{n_{11}}(y_{11k}-\mu_{11})^2 +$$

$$\sum_{k=1}^{n_{12}}(y_{12k}-(p\mu_{11}+(1-p)\mu_{22}))^2 +$$

$$\sum_{k=1}^{n_{21}}(y_{21k}-(p\mu_{22}+(1-p)\mu_{11}))^2 +$$

$$\sum_{k=1}^{n_{22}}(y_{22k}-\mu_{22})^2] \qquad (A.3)$$

where n_{ij} is the number of observations in the ijth cell of the mobility table.

Differentiating (A.3) with respect to μ_{11}, μ_{22}, and p, equating the partials to 0, and rearranging the equations (the equation for σ^2 may be ignored for purposes of the exposition) yields the system:

$$\mu_{11} = (n_{11}\bar{y}_{11} + pn_{12}\bar{y}_{12} -$$

$$p(1-p)(n_{12}+n_{21})\mu_{22} +$$

$$n_{21}(1-p)\bar{y}_{21}) \times$$

$$(n_{11}+p^2n_{12}+(1-p)^2n_{21})^{-1} \qquad (A.4)$$

$$\mu_{22} = (n_{22}\bar{y}_{22} + (1-p)n_{12}\bar{y}_{12} -$$

$$p(1-p)(n_{12}+n_{21})\mu_{11} .$$

[12] Donald Treiman, an anonymous referee, and Goldthorpe (1980) have made this argument.

$$+ n_{21}p\bar{y}_{21}) \text{ X}$$

$$(n_{22}+p^2 n_{21}+(1-p)^2 n_{12})^{-1} \quad (A.5)$$

$$p = (n_{12}(\bar{y}_{12}-\mu_{22})+n_{21}(\mu_{11}-\bar{y}_{21}))\text{X}$$

$$((n_{21}+n_{12})(\mu_{11}-\mu_{22}))^{-1} \quad (A.6)$$

where $\bar{y}_{ij}$ is the sample mean in the ijth cell.

An excessive amount of algebra yields the closed form solution:

$$\hat{\mu}_{11} = (\bar{y}_{11}(n_{22}^{-1} + n_{12}^{-1} + n_{21}^{-1}) +$$

$$n_{11}^{-1} (-\bar{y}_{22} + \bar{y}_{12} + \bar{y}_{21})) \text{ X}$$

$$(n_{12}^{-1} + n_{21}^{-1} + n_{11}^{-1} + n_{22}^{-1}) \quad (A.7)$$

$$\hat{\mu}_{22} = (\bar{y}_{22} \; (n_{11}^{-1} + n_{12}^{-1} + n_{21}^{-1}) +$$

$$n_{22}^{-1} \; (-\bar{y}_{11} + \bar{y}_{12} + \bar{y}_{21})) \text{ X}$$

$$(n_{12}^{-1} + n_{21}^{-1} + n_{11}^{-1} + n_{22}^{-1}) \quad (A.8)$$

$$\hat{p} = [(n_{21}^{-1}\bar{y}_{12}+n_{12}^{-1}\bar{y}_{21})(n_{11}^{-1}+n_{12}^{-1}+n_{21}^{-1}+n_{22}^{-1})$$

$$- n_{21}^{-1}(\bar{y}_{22}(n_{11}^{-1}+n_{12}^{-1}+n_{21}^{-1})$$

$$+ n_{22}^{-1}(\bar{y}_{12}+\bar{y}_{21}-\bar{y}_{11}))$$

$$+ n_{12}^{-1}(\bar{y}_{11}(n_{22}^{-1}+n_{12}^{-1}+n_{21}^{-1})$$

$$+ n_{11}^{-1} \; (\bar{y}_{12}+\bar{y}_{21}-\bar{y}_{22}))] \text{ X}$$

$$[\bar{y}_{11}(2n_{22}^{-1}+n_{12}^{-1}+n_{21}^{-1})$$

$$- \bar{y}_{22}(2n_{11}^{-1}+n_{12}^{-1}+n_{21}^{-1})$$

$$+(\bar{y}_{12}+\bar{y}_{21})(n_{11}^{-1}-n_{22}^{-1})]^{-1} \quad (A.9)$$

This suffices to show the result.

When the number of origin and destinations exceeds 2, closed-form solutions are not readily obtained, and the maximum likelihood equations must be solved by an iterative procedure. As stated in the text, maximum likelihood is equivalent to nonlinear least squares for the problem under consideration. To compute parameter estimates for the models considered in this paper, I differentiated the log likelihood function with respect to these parameters (the p's, the μ's, and the mobility parameters) and set the partial derivatives to 0, obtaining the first-order conditions. Then I input the data and the first-order conditions into a nonlinear least squares program to obtain the solution. I did not use a derivative-free nonlinear least squares program, but some researchers who wish to use the diagonal models may find it more convenient to use a program that does not require inputting the first-order conditions.

It was also necessary to choose starting values for the parameters. For the μ_{ii}, I chose the diagonal means for starting values. For the origin and destination weight parameters, I used the weights obtained by estimating the model by ordinary least squares (OLS), using the diagonal sample means as estimates of the diagonal population means. Similarly, for the mobility parameters, I used the OLS estimates.

It was not necessary, with these data, to directly impose the restriction that the origin and destination weights lie in the closed unit interval. Had it been necessary to do so, the computational procedures are easily modified to handle this situation.

REFERENCES

Baltzell, E. Diby
1953 "Social mobility and fertility within an elite group." Milbank Memorial Fund Quarterly 31:411–20.

Berent, Jerzy
1952 "Fertility and social mobility." Population Studies 5:244–60.

Blau, Peter M.
1956 "Social mobility and interpersonal relations." American Sociological Review 21:290–5.

Blau, Peter M. and Otis D. Duncan
1967 The American Occupational Structure. New York: Wiley.

Boyd, Monica
1971 "Occupational mobility and fertility in Latin America." Demography 10:1–17.

Dumont, Arsène
1890 Dépopulation et Civilisation. Paris.

Duncan, Otis D.
1966 "Methodological issues in the analysis of social mobility." Pp. 51–97 in Neil J. Smelser and Seymour Martin Lipset (eds.), Social Structure and Mobility in Economic Development. Chicago: Aldine.

Gallant, A. R.
1975 "Nonlinear regression." The American Statistician 29:73–81.

Goldfeld, Stephen M. and Richard E. Quandt
1972 Nonlinear Methods in Econometrics. Amsterdam: North-Holland.

Goldthorpe, John H.
1980 "Social Mobility and Class Structure in Modern Britain. Oxford: Clarendon.
Goodman, Leo A.
1972 "Some multiplicative models for the analysis of cross-classified data." Pp. 646–96 in Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability. Berkeley: University of California Press.
Halaby, C. N. and Michael E. Sobel
1979 "Mobility effects in the workplace." American Journal of Sociology 85:385–416.
Hodge, Robert W. and Donald J. Treiman
1966 "Occupational mobility and attitudes toward negroes." American Sociological Review 31:93–102.
Hope, Keith
1971 "Social mobility and fertility." American Sociological Review 36:1019–32.
1975 "Models of status inconsistency and social mobility effects." American Sociological Review 40:322–43.
1981 "Vertical mobility in Britain: a structured analysis." Sociology 15:19–55.
House, James S.
1978 "Facets and flaws of Hope's diamond model." American Sociological Review 43:439–42.
Jackman, Mary R.
1972 "The political orientation of the socially mobile in Italy: a re-examination." American Sociological Review 37:213–22.
Jackson, E. F. and Richard F. Curtis
1972 "Effects of vertical mobility and status inconsistency: a body of negative evidence." American Sociological Review 37:701–13.
Kessin, Kenneth
1971 "Social and psychological consequences of intergenerational occupational mobility." American Journal of Sociology 77:19–35.
Kessler, Ronald C. and Paul D. Cleary
1980 "Social class and psychological distress." American Sociological Review 45:463–78.
Knoke, David
1973 "Intergenerational occupational mobility and the political party preferences of American men." American Journal of Sociology 78:1448–68.
Laslett, Barbara
1971 "Mobility and work satisfaction: a discussion of the use and interpretation of mobility models." American Journal of Sociology 77:19–35.
Malinvaud, E.
1970 Statistical Methods of Econometrics. Amsterdam: North-Holland.
Rao, C. R.
1973 Linear Statistical Inference and its Applications. New York: Wiley.
Riemer, Ruth and Clyde V. Kiser
1954 "Economic tension and social mobility in relation to fertility planning and size of planned family." Milbank Memorial Fund Quarterly 32:167–231.
Rothenberg, Thomas J.
1973 Efficient Estimation with A Priori Information. New Haven: Yale University Press.
Sorokin, Pitrim A.
1927 Social Mobility. New York: Harper.
Stone, Robert C.
1952 "Mobility factors as they affect workers' attitudes and conduct toward incentive systems." American Sociological Review 17:58–64.
Theil, Henri
1971 Principles of Econometrics. New York: Wiley.
Tien, H. Yuan
1961 "The social mobility/fertility hypothesis reconsidered: an empirical study." American Sociological Review 26:247–57.
Westoff, Charles F.
1953 "The changing focus of differential fertility research: the social mobility hypothesis." Milbank Memorial Fund Quarterly 31:24–38.

THE HISTORICAL DISTRIBUTION OF COUNTY SEATS IN THE UNITED STATES: A REVIEW, CRITIQUE, AND TEST OF TIME-MINIMIZATION THEORY*

G. EDWARD STEPHAN
DOUGLAS R. MCMULLIN
Western Washington University

The inverse relation between the size and density of territorial subdivisions has been derived from a general theory which assumes that social structures evolve so as to minimize time expenditure. This paper reviews the empirical work done on the size-density relation, its derivation from time-minimization theory, and the application of that theory to other empirical relations. We derived a new relation: service-center density is proportional to population density raised to the two-thirds power. Unlike the earlier size-density relation, this new relation makes possible a test of the theory even when the areas of subdivisions are unknown, as in the case of United States counties during most of the period of their formation. Finally, we propose an alternative explanation, that random data alone would produce results equivalent to those expected from time-minimization theory. The findings support time-minimization theory but not the alternative explanation.

Counties in the United States vary greatly in size, from the 20,131 square miles of San Bernardino County in southern California, to the 22 square miles of New York County, lying totally within the city of that name. The "theory of segmental growth" (Stephan, 1971) was developed to account for such variation. Based on the prior work of Durkheim (1933), Boulding (1968), and Stinchcombe (1968), the theory argued that increased settlement intensity in a region would lead to its further subdivision, assuming the technology of transportation and communication remained constant. If that were true, large counties should appear in regions of light settlment, small ones in heavily settled regions. Examination of U.S. county outline maps and maps of population density for 1790 and 1840–1970 confirmed this expectation, though the relation appeared to weaken from about 1930. Statistical analysis was not possible because county area data were not available for most of the period. "Visual correlations" only were reported. The essentially verbal theory and qualitative observation procedures are described in detail in the original paper.

EMPIRICAL SUPPORT

Replicative cross-national studies of the relation could not be done because suitable historical maps were not available. The theory was tested, however, with contemporary statistical data from other countries. If a subdivision process had occurred, then we should observe an inverse relation between size and density. Data were obtained for the 1764 primary political divisions (e.g., states, departments, regions) of 98 modern nations. For convenience in plotting area and density values, which vary enormously in range, data were converted to logarithms and subjected to least-squares analysis in the form

$$\log A = K + b(\log D) \qquad (1)$$

where A is area, D is density, K is the intercept, and b is the slope, which, according to theory, should be negative.

The slope was negative in all but 4 of the 98 cases. When aggregated, the set of 1,764 subdivisions gave the "world regression line":

* Direct all correspondence to G. Edward Stephan, Department of Sociology, Western Washington University, Bellingham, Washington 98225.

$$\log A = K - 2/3(\log D) \qquad (2)$$

These results were reported in Stephan (1972).

Cross-cultural studies provided still further support for the size-density relation. Territorial units were studied at the time of Western contact for the aboriginal peoples of the Pacific Northwest (Stephan and Wright, 1973); Africa (Stephan and Tedrow, 1974); Melanesia (Callen and Stephan, 1975); and California (Myers and Stephan, 1976). Other researchers found negative size-density relations for subdivisions in Brazil (Haggett, 1965), mainland China (Skinner, 1964), and New Zealand (Webb, 1974).

Some unpublished work has suggested that population-specific densities may give a better fit than can be obtained from general population densities. Hall (1973) found that police patrol districts in the city of Seattle showed no relation between size and population density but did show negative relations between size and "crime density" (reported patrolable crimes per square mile). The relation held even though the district boundaries shifted three times each day with a change of watch. Suggs (1977) found that diocesan boundaries of the Episcopal and Roman Catholic churches in the United States, though conforming to the size-density hypothesis, produced still stronger correlations when the independent variable was "denominational density" rather than general population density.

Analysis of a deviant case provided some further evidence for the negative two-thirds slope of Equation 2. Massey and Stephan (1977) investigated the relationship for Great Britain from 1801 to the present, spurred by the fact that the relationship for the case had been just barely negative in an earlier study (Stephan, 1972). With each census going back to the earliest, the slope more nearly approached −2/3, suggesting that the relation held before the Industrial Revolution and decayed as population shifted while boundaries remained fixed. In 1973, when for the first time in a thousand years the British counties were drastically reconstituted, a slope of −2/3 was generated. This slope was subsequently reconfirmed using regional rather than local densities (Massey, Tedrow, and Stephan, 1980).

TIME-MINIMIZATION

In addition to the Durkheim-Boulding-Stinchcombe model used in Stephan (1971) to account for county size variations, several other theoretical accounts have been suggested. Fox (1976) argued that the carrying capacity of the land determines density and, hence, the relation between size and density. Stephan (1976) disputed this, but noted an earlier explanation (Stephan and Wright, 1973) in which carrying capacity may play an important role under certain circumstances. In the same earlier study additional geographical, cognitive and ecological explanations were suggested. Schuessler (1974) noted that the size-density relation may be the result of arbitrary political action, official regulations placing limits to area and/or population of subdivisions. This was not in fact the case during most of the period of county subdivision in the United States and could hardly be said to apply cross-nationally or cross-culturally where there is no superordinate governing system.

All these theoretical explanations only account for the negativity of the relation, not the negative two-thirds value given in Equation 2, which was, however, derived from time-minimization theory (Stephan, 1977). One factor common to all previous theoretical explanations was distance, as a cost to be overcome in spatial interaction. If territories are too large, the distances within them make significant interaction impossible. If territories are too small, they may not have the natural or human resources needed to maintain them. Both distance and maintenance requirements can be expressed in terms of time-expenditure by the population of the territorial unit, providing an expression for the average time expenditure:

$$T = s/v + h/p \qquad (3)$$

where s is the average distance to be traversed, v is the average velocity of transportation, and h is the total maintenance time extracted from the population

p. Since distance must be proportional to the square root of area (so $s = g\,a^{1/2}$, where g is the constant of proportionality), and since density equals population divided by area (so $p = ad$), the two terms in Equation 3 can be rewritten

$$T = ga^{1/2}/v + h/ad$$

By taking the derivative of T with respect to a,[1] one determines the condition under which time is minimized,

$$a = cd^{-2/3} \quad (4)$$

the logarithmic form of which is simply Equation 2.

Virirakis (1969), beginning with somewhat different assumptions, concludes, for ideal circular areas only, that population-times-circular-radius should be constant. This result is algebraically equivalent to Equation 4 but does not provide an appropriate statistical test of the theory (cf. Vining et al., 1979; Stephan, 1979c; and below).

In Stephan and Tedrow (1977) equations were derived from time-minimization theory to relate area and population for both central cities and urbanized areas. Data from the 1950, 1960, and 1970 censuses confirmed these equations. Stephan (1979) also derived the interactance or gravity model, the urban density function, and rank-size rule from the theory. Finally, time-minimization has been used to account for a variety of univariate, continuous probability density functions (Stephan, 1979b). Each empirical relationship is derived from the same general assumption: social structures evolve in such a way as to minimize the time expended in their operation.

[1] The first derivative is

$$dT/da = g/2va^{1/2} - h/a^2d$$

which, set equal to zero, gives

$$a^{3/2} = 2vh/gd$$

so that

$$a = cd^{-2/3}$$

where $c = (2vh/g)^{2/3}$. The second derivative is greater than zero, so Equation 4 shows the relation between a and d when T is a minimum.

PROBLEM STATEMENT

Time-minimization theory appears to account for a wide variety of different empirical findings, and these findings have themselves received extensive empirical support in a wide variety of settings. However, the size-density law in its final form (Equation 4) has not been tested where it in a sense was born: the historical process of county subdivision in the United States. It would seem useful to test that equation historically and statistically for the United States; however, we remain confronted by the absence of relevant areal data for much of the historical period of county formation. Given this state of affairs, we must derive an alternative to Equation 4 which can be tested without areal data. If Equation 4 is true, then the expected area of a county E(a) in a region will be given by inserting the expected density E(d) for that region into Equation 4. The expected density is simply the average density throughout the region, and that will be the average of county densities weighted by area, $\Sigma a_i d_i / \Sigma a_i$, or, since $d_i = p_i/a_i$, the expected density is $\Sigma p_i / \Sigma a_i$. If we let P stand for Σp_i and A stand for Σa_i, the expected density is $D = P/A$, the overall density of the region, and we have

$$E(A) = cD^{-2/3} \quad (5)$$

We obtain another expression for E(a), the expected area of a county, from the usual definition of expectation in statistics, simply the mean or average, A/N, where N is the number of counties in the region. We equate our two expressions for E(a), and, by simple algebra, obtain

$$C = KD^{2/3} \quad (6)$$

where $C = N/A$ and the constant K is simply 1/c. Equation 6 says that the density of counties (county seats per square mile) will be proportional to the population density raised to the two-thirds power. The expression is identical to one derived from geographic considerations by Palmer (1973), and to a still earlier one

derived from economic considerations by Mycielski and Trzeciakowski (1963). In both works the derivations are considerably more complex and the assumptions are more restrictive than in the derivation from time-minimization. That is, time-minimization theory is both simpler and more generally applicable than are the earlier derivations.

METHODOLOGY

We now have an expression which, in its logarithmic form,

$$\log C = \log K + b(\log D) \quad (7)$$

can be tested for some suitable region. Issues yet to be resolved are (1) specifying the regions or units to be studied; (2) determining whether C and D, as ratio variables, are proper variables for the test of the theory of time-minimization, since each contains the common denominator A and the relation may thus be spurious; (3) determiniation of the method of estimating b empirically (i.e., is least-squares linear regression suitable for the logarithmically transformed variables?); and (4), assuming that the two-thirds value in Equation 7 is obtainable, determination of whether any conditions other than those specified by time-minimization theory could produce such a result.

Units of Analysis

Two types of regions may be used as units of analysis: geographic quadrants defined by degrees of latitude and longitude and independent of county areas themselves, and states as "natural" collections of counties. Quadrants have certain advantages: they are relatively small units so that the average or expected density in Equation 7 can more reasonably be expected to approximate conditions throughout the region. Unfortunately, there are also disadvantages. In order to obtain the county-density of a quadrant we need only determine the number of county seats there and divide by area. But to obtain population density we must determine the population of the quadrant. This could be done by assigning the entire population of a county to the quadrant where its seat is located, but this procedure would arbitrarily increase the correlation implied by Equation 7, especially for larger counties. That is, the methodologically induced arbitrariness of the results would increase in lightly settled regions and during earlier time periods when counties were larger.

We select states (or equivalent territorial entities, see below) as units of analysis for several reasons. First, the resulting data are not ambiguous or subject to arbitrary manipulation: a state's county seats, area, and population are all clearly assignable to that state. Second, selecting the state as a unit actually works against confirmation of the hypothesis: Equation 7 clearly depends on the degree to which one can reasonably assume uniform or average density throughout the region. Instead of the internally homogeneous quadrants, we have much larger states, where the assumption is certainly less likely to hold. Finally, we select states simply because they are the "natural" areal units for sets of county areas (Duncan, Cuzzort, and Duncan, 1961:50); counties are creatures of state governments and exist wholly within their boundaries.

Ratio Variables

A number of writers (Freeman and Kronenfeld, 1973; Fuguitt and Lieberson, 1974; Schuessler, 1974) have argued that expressions such as Equation 7 may be inappropriate since they contain ratio variables with common terms.[2] Since C and D have A as a common denominator, any correlation observed would be spurious. Vining et al. (1979) objected to Stephan's test of Equation 1, stating that "it merely confirms an artifactual relation"

[2] Snedecor (1946:162) pointed out that, from the point of view of statistical methodology, correlating ratios is no different from any other kind of correlation. Stephan (1972) called attention to this issue in his own research; but, in spite of his discussion of the matter, subsequent writers (Fox, 1976; Vining et al., 1979) have continued to describe the size-density relationship as "spurious" or as a "statistical artifact." Fortunately, a fairly recent and thorough discussion of this issue should eliminate the need for extensive consideration of it here (MacMillan and Daft, 1980; Kasarda and Nolan, 1979).

between a and p/a. He proposed instead to decompose the relation, obtaining an algebraic equivalent in the form of a relation between a and p. By a similar procedure, we could easily derive

$$N = KP^{2/3}A^{1/3} \qquad (8)$$

from Equation 6—by multiplying through by A—and test its logarithmic form using multivariate regression techniques.

The choice here is not a trivial one, since algebraic equivalence between the two expressions does not imply empirical or statistical equivalence between them. Except under conditions of perfect correlation, the regression coefficient of b_{xy} does not even imply its own inverse b_{yx}. Confirmation or rejection of one of the two equations above does not imply confirmation or rejection of the other. Clearly, we must choose one or the other as the more appropriate test of the theory.

Equation 7 is the appropriate test of the theory since it is the relative spatial distribution of P and N, not the relation between their absolute sizes, which is predicted from considerations of time-minimization. It is the spatial distribution of a population, not its size, which makes territorial subdivision necessary; and population density is a measure of this distribution. In Equation 4 the dependent variable is the area of a county; at the state level, the appropriate measure must be the effect of the component areas, i.e., the overall density of county seats, not simply their absolute number. Far from representing a spurious relationship, Equation 7 actually controls for the common element A by standardizing P and N in terms of it. Other things equal, large areas contain large numbers of many things: trees, rocks, people, and county seats. If so, it would be no trick to correlate N and P, but such a relation would then surely be spurious, at least as that word is commonly employed. The argument is the same as that which pertains to correlating suicide and employment rates rather than suicides and number unemployed, regardless of population size.

Logarithmic Transformation

Is the logarithmic transformation of Equation 6 appropriate? That is, can we justify applying regression techniques to Equation 7? Research conducted on untransformed variables seldom concerns itself with the normality assumptions behind regression analysis; the normal curve is presumed to be so prevalent in nature that it is simply assumed in most regression studies. Such tacit acceptance tends to be questioned when variables are transformed from their "natural" state. As Hays and Winkler have noted: "Minimizing the sum of squared deviations for the transformed model is not necessarily equivalent to minimizing the sum of squared deviations for the original model" (1971:653). The difficulty lies in the transformation of the error term, and it is not automatically a violation of procedure to make such a transformation since "it is possible the original terms violate these assumptions while the transformed terms do not" (1971:654).

Contrary to the tacit assumption of normality in most research, normal distributions are not particularly "normal" in most sociological variables. As Aitchison and Brown (1963:XI) have demonstrated and as Pareto (1966:130–1) noted eighty years ago, most variables in sociology and economics are highly skewed. If the distribution of a variable is skewed, then it may be necessary to transform the variable in some way to produce a normal distribution suitable for regression analysis. We argue that the variables C and D in Equation 7 are lognormally distributed and hence can only be properly analyzed through logarithmic transformation.

A variate is lognormally distributed if the logarithm of that variate is normally distributed. Such distributions can result from the operation of the "law of proportionate effect," which holds whenever the change in a variate at any stage of a process is a random proportion of the previous value of the variate (Aitchison and Brown, 1963:22). If the rich get richer, if early learning affects later learning, if nothing succeeds like success, then the law of proportionate effect holds and the resulting distribution is lognormal. We suggest that the component variables N, P, and A are subject to this law and that

the composite variables C and D are thus lognormal.[3]

We believe both N and P, at least during the period of county formation, were subject to the law of proportionate effect. Regions which originally attracted only small populations tended to grow slowly with few county seats being established. The settlers who occupied more attractive regions, on the other hand, drew still more settlers, leading to the creation of still more county seats to serve them. Of course there were upper limits to such increases, but that is not a problem, since all that is required for lognormality is that the law of proportionate effect operate during the period of growth (Cramér, 1946). Finally, with respect to the variable A, there is another theoretical model which generates lognormal distributions, the so-called "theory of breakage" (Aitchison and Brown, 1963:26–7; Bender, 1978:207–12), which is really a modification of the law of proportionate effect. If counties were formed through a process of segmental subdivision of larger original counties—which they were (Stephan, 1971)—then the theory of breakage applies and A is lognormally distributed. If N, P, and A are lognormally distributed, then so are C and D, since, if x_1, x_2, and x_3 are lognormal variates then the ratios x_1/x_3 and x_2/x_3 are lognormal variates (Aitchison and Brown, 1963: theorem 2.3).

In the absence of contradictory evidence, and on the theoretical arguments just presented, we conclude that regression analysis of Equation 7 is justified. These considerations provide still further justification for choosing Equation 7 over Equation 8. Regression analysis assumes no statistical error in the independent variable. Equation 8 adds the error components associated with log P and log A while Equation 7 minimizes the error term since the multiplicative errors of P and A tend to cancel out in D = P/A (or the additive errors tend to subtract in log D = log P − log A).

[3] In spite of the fact that N,P, and A certainly look lognormal when tabulated, or at least highly skewed, our suggestion will have to be purely theoretical since there are simply too few cases for an adequate empirical test.

An Alternative Interpretation

The final criticism of Equation 7 as a test of time-minimization theory is related to the issue of spuriousness dealt with above, but it cannot be resolved by pointing to the theoretical importance of the relevant ratio variables. The problem is that the relationship derived from time-minimization theory can also be derived from the assumption of a completely random process. Equation 7 would yield b = 2/3 if the analysis were conducted on completely random data. To see how this could happen, draw four sets of random numbers and label the first two sets P and N. Now multiply pairs of values from the remaining two sets (random north-south and east-west distances?) and label the set of products A—a legitimate operation since area is two-dimensional while P and N are one-dimensional. It is obvious that the covariances S_{pn}, S_{pa}, and S_{an} (where p = log P, n = log N, and a = log A) will equal zero. And, since log A will be the sum of the logs of its two component random numbers, the variance of log A will be twice that of log P.

Now consider what happens when we test Equation 7 with such data. The standard formula for the regression coefficient is

$$(n\Sigma xy - \Sigma x\Sigma y)/\ n\Sigma x^2 - (\Sigma x)^2$$

where, in our case, x = log D, or log P − log A, and y = log C, or log N − log A. Making these substitutions, it can be shown that the regression coefficient will be

$$(S_{pn} - S_{pa} - S_{an} + S_{aa})/$$
$$(S_{pp} + S_{aa} - 2S_{pa})$$

where S_{pp} is the variance of Log P and S_{aa} is the variance of log A. Since the covariances are zero and since the variance of log A is two times that of log P, the result is that b = 2/3 whether we assume time-minimization or simple randomness.

If we obtain the same result either way, who needs time-minimization theory (whatever its utility in previous derivations)? Can the principle of time-minimization add anything to what is already known—or rather not known—

under conditions of pure randomness? First of all, the situation is not novel. As is well illustrated in the field of statistical thermodynamics, under conditions of complete equilibrium, random processes in nature result in least-energy outcomes described by quite specific equations (Andrews, 1975; Bent, 1965). To the extent that the expenditure of time to accomplish activities in social systems parallels the expenditure of energy in performance of work in physical systems—or to the extent that statements of time-energy equivalence are conceivable—to that extent the fact that a random process can produce results compatible with time-minimization theory is not, or should not be, surprising.

Furthermore, to the extent that one conceives of a social system as economists conceive of a pure market—a myriad of private activities taking place for private motives, but such that no one individual or group significantly structures the behavior of the aggregate—to that extent one can legitimately treat the macro-level parameters of the system as the result of a random process. If, whatever their individual motives, the members of a population share the common denominator of minimizing the time they devote to certain necessary (but not particularly desired) activities—such as occasional interaction with a governmental center—then the time-minimization assumption is justified even if an alternative assumption of randomness would lead to similar results.

The issue need not remain at this level of purely scholastic debate, however. In the end there are empirical answers to the question of whether an observed result was effected by randomization and/or time-minimization. We have but to examine the data to see if the conditions of randomness are met. Are the relevant covariances equal to zero, and is the variance of log A twice that of log P?

REPORT OF FINDINGS

Data were obtained from two sources: we obtained the number of counties in each unit for each census year from Stephan (1971:453), and the area and population figures from a standard reference (United States, 1975:Series A 195–209 and Series A 210–63). "Population" in this study is "rural population" because the theory does not address itself to urban concentrations but rather to the dispersed population for whom counties represented the relevant units of local government. The populations comprising the 1764 political subdivisions represented by the "world regression line," Equation 2, were also overwhelmingly rural. With these data for N, P, and A, regression analyses were conducted for each census year from 1790 through 1970. The results of the analyses are shown in the accompanying table and figures.

The coefficient of determination r^2 is quite high throughout the period, with explained variance ranging from 72 to 90% (see Table 1). The regression coefficient b rides a roller-coaster, first hovering near the theoretical value of 2/3, sliding downward, rising in 1850, then back down slightly, then upwards toward the theoretical value, finally sliding steadily downward since 1910 (see Figure 1). This behavior is not as erratic as it might at first seem and, in fact, provides a more detailed picture of the growth of county government than Stephan (1971) gave from simple comparison of maps for the overall period. The figure begins in 1790 when the nation is in a state of relative territorial and demographic equilibrium, with territorial and demographic growth having reached some stability by the end of the colonial and confederationist periods. From this point on, under the doctrine of Manifest Destiny, new territories are acquired, gradually populated, and only then given county seats. The frontier areas are at first relatively undergoverned, and b is appropriately lower than the expected 2/3. What governments they do have are costly in both travel and maintenance time. Total time is not minimized in these regions until the predicted density of county seats is finally brought about near the end of the century. The slight increase in b with the census of 1850 reflects the admission of two important units to the Union following the war with Mexico, namely the former Republics of Texas and California which brought with them 80 and

Table 1. Summary of Findings (See text for explanation of symbols.)

Census Year	No. of Units	Covariances p,n	Covariances a,n	Covariances p,a	Variances a	Variances p	Variances a:p	Test of Hypotheses r^2	Test of Hypotheses b	Test of Hypotheses p(b = 2/3)
1790	18	.54	.55	.22	1.60	.73	2.2	.85	.72	.46
1800	22	1.51	.27	−.23	1.65	2.90	.6	.92	.66	.84
1810	27	1.63	.12	−.55	1.40	2.90	.5	.91	.64	.56
1820	27	.70	.51	−.08	1.41	1.61	.9	.86	.57	.035
1830	28	.56	.75	.11	1.33	1.16	1.2	.72	.46	.00082
1840	30	.62	.80	.21	1.31	1.14	1.1	.72	.46	.00052
1850	36	.97	.44	−.39	1.63	1.97	.8	.86	.58	.026
1860	40	.97	.63	−.13	1.42	1.96	.7	.82	.52	.00048
1870	46	1.31	.45	−.30	1.27	2.62	.5	.90	.54	.000045
1880	47	1.14	.53	−.02	1.23	2.00	.6	.89	.60	.036
1890	48	.96	.59	.17	1.20	1.41	.9	.86	.62	.22
1900	48	.92	.62	.25	1.20	1.34	.9	.89	.62	.16
1910	48	.84	.65	.34	1.20	1.08	1.1	.87	.65	.71
1920	48	.80	.70	.41	1.20	1.02	1.2	.86	.64	.43
1930	48	.78	.70	.41	1.20	1.00	1.2	.82	.63	.33
1940	48	.75	.70	.40	1.20	.99	1.2	.80	.61	.20
1950	48	.69	.70	.31	1.20	.88	1.4	.80	.60	.15
1960	48	.63	.70	.27	1.20	.83	1.4	.76	.57	.056
1970	48	.60	.70	.21	1.20	.82	1.5	.74	.55	.023

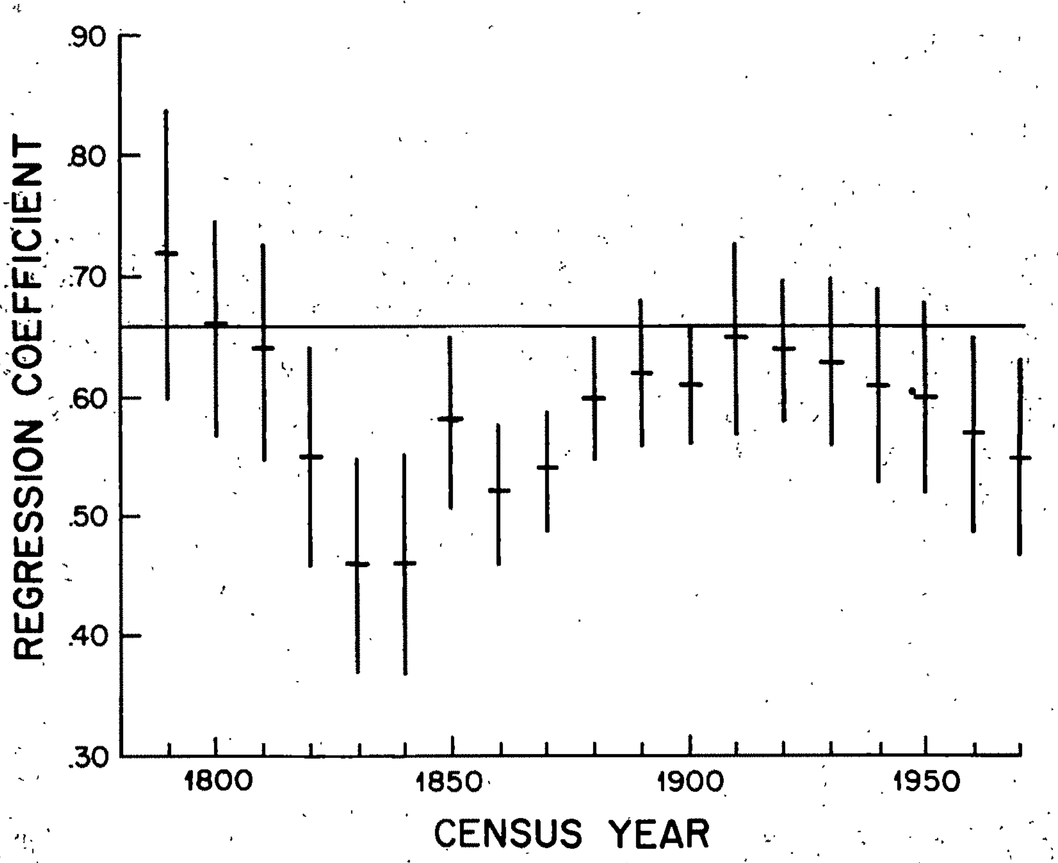


Figure 1. Regression Coefficients and their 95% Confidence Intervals (Vertical Lines), 1790–1970 (The horizontal line is the theoretically expected slope.)

27 units, respectively, of previously established Mexican local government, units which were simply redesignated counties upon admission. These areas had been locally governed by county-like units for more than a century. After the Indian wars, with still more settlement during the remainder of the century, more counties became established. With the closure of the frontier in 1890 (Turner, 1962), settled regions have acquired sufficient county seats to again bring b within range of the theoretically expected value of 2/3.

The theoretical equilibrium value thus obtained begins to erode after 1910. The reason why it should has already been suggested (Stephan, 1971): the introduction of the automobile (1 vehicle registration in 10,000 people in 1900, compared to 1 in 5 by 1930, cf. Abrahamson, 1980:306) made it unnecessary to create new county seats as people moved into what had been frontier areas under horse-and-buggy technology. While hardly resulting in "effacement of the segmental type," as predicted by Durkheim (1933:181), segmental growth had at least ceased in the face of continued expansion of settlement (Figure 1). Whether one can apply theory or the results of previous observation, as in the case of Great Britain, to predict some future restructuring of boundaries, might be a good topic for future work in applied sociology but cannot be pursued here.

From Table 1 and Figure 2 it is evident that these values of b are not the result of simple randomness. The covariances may seem small (that is, near zero) for those unaccustomed to working with logarithms. But note that the variances for logarithms of area are not great though the areas range from that of Delaware to that of Texas. The only covariance close to zero is S_{pa}, and its approach toward that value does not correspond with the approach of b toward 2/3.

The ratio of the variance of log A to that of log P is not on the order of 2:1 except in 1790 (Table 1). As it begins to approach that value since 1910, b departs steadily from its 2/3 value. There is no consistent pattern between the value of b in Figure 1 and the b-component values shown in Figure 2. The three distinguishable sections of Figure 2—settlement of the Midwest through 1840, settlement of the far west through 1890, and then the twentieth century—do not show variations in the components of b which coordinate routinely with the decline, rise, and decline which Figure 1 shows for b. So the random process suggested earlier does not

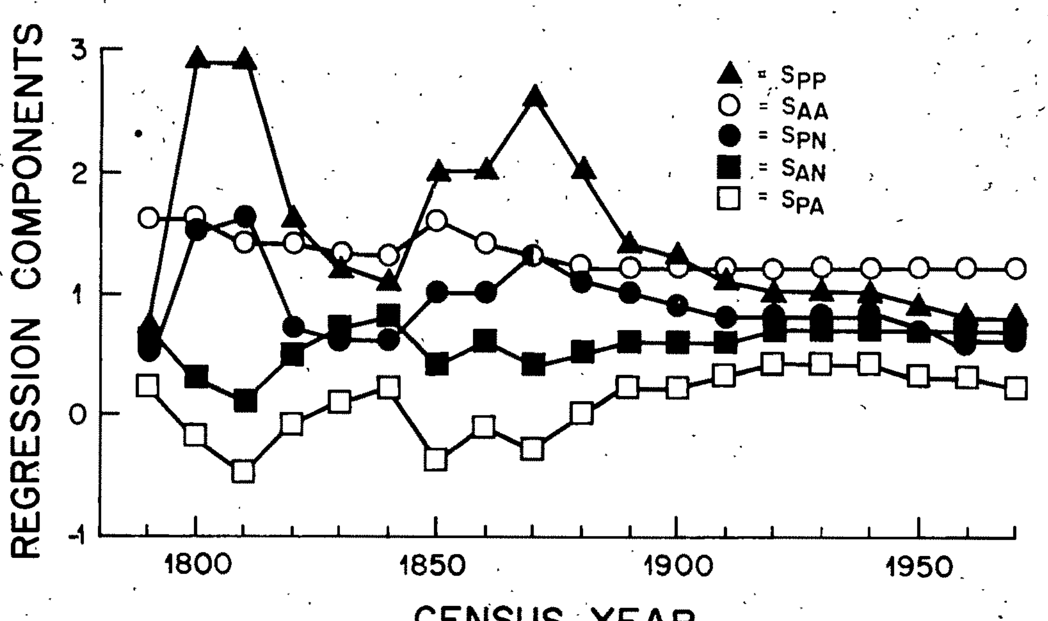


Figure 2. Variances and Covariances for log P, log A, and log N, 1790–1970

appear to account empirically for the 2/3 value observed for Equation 7. We are left only with the theory of time-minimization.

CONCLUSION

We conclude that the theory of time-minimization has been supported by the data. Equation 7, which is derived from that theory, is neither spurious, nor a statistical artifact, nor an inappropriate form to test, nor simply the result of a random process. Rather, it accurately describes the interaction between two populations: a human population and a "social population" of governmental units which touched those people closely during much of their history. More generally, Equation 7 should apply to the spatial distribution of any population of social structures (schools, churches, retailing firms), so long as those structures must interact with a distributed population (students, church goers, consumers), and provided the assumptions of time-minimization theory apply.

Finally, since time-minimization theory has already been applied to derive a diverse set of empirical relations—beyond the size-density one which gave rise to it—we suggest future efforts to derive non-ecological relations (for example, empirical relations from organizational research) from the theory. If such efforts prove successful, time-minimization theory could become as useful to sociology as the principle of energy-minimization (entropy-maximization) has been to the physical scientist or the principle of cost-minimization (benefit-maximization) has been to the economist—enabling us to derive precise empirical relations while providing a coherent link between our diverse research areas.

REFERENCES

Aitchison, J. and J. A. C. Brown
1963 The Lognormal Distribution. Cambridge: University Press.

Abrahamson, Mark
1980 Urban Sociology, Second Edition. Englewood Cliffs: Prentice Hall.

Andrews, Frank C.
1975 Equilibrium Statistical Mechanics, Second Edition. New York: Wiley.

Bender, Edward A.
1978 An Introduction to Mathematical Modeling. New York: Wiley.

Bent, Henry A.
1965 The Second Law: An Introduction to Classical and Statistical Thermodynamics. New York: Oxford.

Boulding, Kenneth
1968 "The city as an element in the international system." Daedalus 97:1111–23.

Callen, Jay S. and G. Edward Stephan
1975 "Siwai line-villages: Thiessen polygons and the size-density hypothesis." Solomon Island Studies in Human Biogeography, Occasional Paper No. 4, Field Museum of Natural History, Chicago.

Cramér, H.
1946 Mathematical Methods of Statistics. Princeton Mathematical Series, no. 9. Princeton: University Press.

Duncan, Otis Dudley, Ray P. Cuzzort, and Beverly Duncan
1961 Statistical Geography: Problems in Analyzing Areal Data. Glencoe: Free Press.

Durkheim, Emile
1933 [1893] The Division of Labor in Society. Translated by George Simpson. Glencoe: Free Press.

Fox, John G.
1976 "Tests of the size-density hypothesis: a caution." American Sociological Review 41:567–9.

Freeman, J. H. and J. E. Kronefeld
1973 "Problems of definitional dependency: the case of administrative intensity." Social Forces 52:108–21.

Fuguitt, G. V. and S. Liebersen
1974 "Correlation of ratios or difference scores having common terms." In Herbert L. Costner (ed.), Sociological Methodology 1973–1974. San Francisco: Jossey-Bass.

Haggett, Peter
1965 Locational Analysis in Human Geography. London: Edward Arnold.

Hall, Greg
1973 "Administrative areas of the Seattle police department: an application of the size-density hypothesis." Master's Thesis, Department of Sociology, Western Washington University.

Hays, William L. and Robert L. Winkler
1971 Statistics: Probability, Inference, and Decision. New York: Holt, Rinehart and Winston.

Kasarda, John D. and Patrick D. Nolan
1979 "Ratio measurement and theoretical inference in social research." Social Forces 58:212–27.

MacMillan, Alexander and Richard L. Daft
1980 "Relationships among ratio variables with common components: fact or artifact." Social Forces 58:1109–28.

Massey, Douglas S. and G. Edward Stephan
1977 "The size-density hypothesis in Great Britain: analysis of a deviant case." Demography 14:351–61.

Massey, Douglas S., Lucky M. Tedrow, and G. Edward Stephan
1980 "Regional population density and county size: a note on the problem of tautology in size-density relationships." Geographical Analysis 12:184–8.

Mycielski, Jerzy and Witold Trzeciakowski
1963 "Optimization of the size and location of service stations." Journal of Regional Science, 5:59–68.

Myers, David E. and G. Edward Stephan
1976 "Tribal territories of the California Indians: a test of the size-density hypothesis." Anthropology UCLA 6:59–66.

Palmer, D. S.
1973 "The placing of service points to minimize travel." Operational Research Quarterly 24:121–3.

Pareto, Vilfredo
1966 [1902] "Less systèmes socialistes." Translated by Derick Mirfin. Pp. 123–43 in S. E. Finer (ed.), Vilfredo Pareto: Sociological Writings. New York: Praeger.

Schuessler, Karl
1974 "Analysis of ratio variables: opportunities and pitfalls." American Journal of Sociology 80:379–96.

Skinner, G. William
1964 "Marketing and social structure in rural China." The Journal of Asian Studies 24:3–43.

Snedecor, George
1946 Statistical Methods, 4th edition. Ames, Iowa: Iowa State College Press.

Stephan, G. Edward
1971 "Variation in county size: a theory of segmental growth." American Sociological Review 36:451–61.
1972 "International tests of the size-density hypothesis." American Sociological Review 37:365–8.
1976 "Reply to Fox." American Sociological Review 41:569–70.
1977 "Territorial division: the least-time constraint behind the formation of subnational boundaries." Science 196:523–4.
1979a "Derivation of some social-demographic regularities from the theory of time-minimization." Social Forces 53:812–23.
1979b "A theory and method for deriving the distribution of sociological variables." Technical Report #04-30-79, Demographic Research Laboratory, Department of Sociology, Western Washington University.
1979c "Reply to Vining et al." Science 205:219–20.

Stephan, G. Edward and Lucky M. Tedrow
1974 "Tribal territories in Africa: a cross-cultural test of the size-density hypothesis." Pacific Sociological Review 17:365–9.
1977 "A theory of time-minimization: the relationship between urban area and population." Pacific Sociological Review 20:105–12.

Stephan, G. Edward and Stephen M. Wright
1973 "Indian tribal territories in the Pacific Northwest: a cross-cultural test of the size-density hypothesis." Annals of Regional Science 7:115–23.

Stinchcombe, Arthur L.
1968 Constructing Social Theories. New York: Harcourt, Brace and World.

Suggs, Glenn W.
1977 "National religious administrative districts: a test of the size-density hypothesis." Master's Thesis, Department of Sociology, Western Washington University.

Turner, Frederick Jackson
1962 [1920] The Frontier in American History. New York: Holt, Rinehart, and Winston.

United States Bureau of the Census
1975 Historical Statistics of the United States, Colonial Times to 1970, Bicentennial Edition. Washington, D.C.: USGPO.

Vining, Daniel R., Jr., Chung-Hsin Yang, and Shi-Tao Yeh
1979 "Political subdivision and population density." Science 205:219.

Virirakis, John
1969 "Minimum effort as a determinant of the area, population and density of residential communities." Ekistics 27:362–71.

Webb, Steven
1974 "Segmental urban growth: some cross-national evidence." Sociology and Social Research 58:387–91.

RACIAL CHARACTERISTICS AND THE IMPOSITION OF THE DEATH PENALTY*

MICHAEL L. RADELET
University of Florida

This paper explores the question of whether death penalty statutes passed after the 1972 Supreme Court decision in Furman v. Georgia successfully eliminate racial disparities in capital cases. Over 600 homicide indictments in twenty Florida counties in 1976 and 1977 were examined, focusing on homicides between strangers (nonprimary homicides). Those accused of murdering whites are more likely to be sentenced to death than those accused of murdering blacks. This trend is due primarily to the higher probability for those accused of murdering whites to be indicted for first degree murder. When controlling for race of the victim, the data do not clearly support the hypothesis that race of the defendant is strongly associated with the probability of a first degree murder indictment or the imposition of the death penalty.

Few aspects of the American criminal justice system have generated as much controversy as the reemergence of the threat and use of capital punishment. Central to this debate is the assertion that the death penalty is disproportionately applied to black offenders. Faced with evidence supporting this claim, when Furman v. Georgia (408 U.S. 238 [1972]) came before the U.S. Supreme Court in 1972, the Court ruled (5 to 4) that the application of all death penalty statuses then in existence was arbitrary and capricious, hence constituting "cruel and unusual punishment" in violation of the Eighth Amendment (Bowers, 1974; Bowers and Pierce, 1980; Riedel, 1976). Since that ruling, 37 states and the federal government have enacted new capital punishment statutes that have not been invalidated by the courts (Gregg v. Georgia, 428 U.S. 153 [1976]; U.S. Department of Justice, 1980). The purpose of this paper is to ascertain whether or not race remains a significant factor in the processing and outcome of post-Furman homicide cases. Data to be examined are based on 637 homicide indictments in twenty Florida counties in the years 1976 and 1977.

There is little doubt that the history of capital punishment in the United States prior to the Furman decision is marked by inequality and discrimination (Bowers, 1974; Garfinkle, 1949; Sellin, 1980). Of the 3,859 civil (i.e., nonmilitary) executions that took place from 1930 through 1967, 2,108, or 54%, involved nonwhite offenders (U.S. Department of Justice, 1980). Blacks were especially overrepresented among those executed for rape: 405 of the 455 executions for this crime, or 89%, were of blacks (U.S. Department of Justice, 1980). One possible explanation for this differential is that although the percentage of blacks executed is higher than their representation in the nation's population, the percentage executed might approximate their conviction rates. However, in an analysis of 1,265 rape convictions in seven southern states between 1945 and 1965, Wolfgang and Riedel (1973) found that 13% of the 823 blacks convicted of rape were sentenced to death, whereas only 2% of the 442 convicted whites confronted a similar fate. Signifi-

* Direct all correspondence to Michael L. Radelet, Department of Sociology, University of Florida, Gainesville, FL 32611.

I would like to thank Ronald L. Akers, Ronald A. Farrell, Pamela Richards, Charles W. Thomas, and especially Alan Agresti and John C. Henretta for their helpful comments on earlier drafts, and Jack Geckler for his assistance in coding the data. The data were collected and analyzed with funds provided by the Edna McConnell Clark Foundation, the North Shore Unitarian Veatch Program, and the NAACP Legal Defense and Educational Fund, Inc. This paper was presented at the 1981 Annual Meeting of the American Sociological Association in Toronto.

cant racial differences remained even after controlling for a variety of nonracial aggravating circumstances, such as the defendant's prior record and the commission of additional felonies concurrently with the rape.

The relationship between race and the imposition of the death penalty is only part of the larger issue of the extent to which racial disparities are evident in the selection, processing, and sentencing decisions of the criminal justice system. Several authors have asserted that violations of the existing criminal law by lower class and minority offenders result in more severe charges and sentences (Black, 1976; Chambliss and Seidman, 1971; Quinney, 1979). Yet, others have failed to find substantial racial differences (Burke and Turk, 1975; Hindelang, 1969, 1978). In a review of seventeen studies that investigated the effects of race on sentencing, Hagen found that most relationships explained little of the sentencing variation. Chiricos and Waldo (1975), in an analysis of 10,488 felony sentences in three southeastern states between 1969 and 1973, found that the socioeconomic status of convicted felons was unrelated to the severity of prison sentences. Similarly, in a sample of 1,213 males arraigned on a felony charge in New York state in the four years preceding March, 1975, Bernstein et al. (1977) found only minor effects of race on the severity of imposed sentences. Clearly, any current tendency for the criminal justice system to treat blacks more harshly than whites is not as apparent or widespread as has been alleged.

Three limitations of these studies are addressed in this paper. First, each focuses only on the characteristics of the defendant, leaving open the possibility that processing decisions might vary in part with such characteristics as the race of the victim. Second, all cases of a crime are uniformly categorized, thus ignoring qualitative differences within the crime category itself that may be associated with sentencing variations (e.g., Garfinkle, 1949; Swigert and Farrell, 1977). Third, with some exceptions (Bernstein et al., 1977; LaFree, 1980; Swigert and Farrell, 1977), many studies of racial disparities in the criminal justice system have focused primarily on sentencing decisions made by judges (Chiricos and Waldo, 1975; Hagen, 1974; Unnever et al., 1980). Racial disparities may also be evident in decisions made at other points in the criminal justice process (Thomson and Zingraff, 1981).

Several studies have suggested that the general lack of data on victim characteristics is a crucial omission if the effects of race on case outcome are to be understood (Myers, 1979). LaFree (1980), for example, in a study of 881 sexual assault suspects in a large midwestern city for the years 1970 through 1975, found strong evidence that black men accused of assaulting white women received more severe sentences than other sexual assault defendants. Similarly, in a study of the records of 821 persons indicted for homicide in ten North Carolina counties during the 1930s, Garfinkle (1949) found that blacks accused of killing whites were significantly more likely to be indicted for first degree murder than were whites accused of killing blacks. In the rape study by Wolfgang and Riedel (1973), 113 of the 317 black defendants with white victims were sentenced to death (36%), whereas only 19 of the 921 defendants (2%) with all other victim-defendant racial combinations were sentenced to death.

A second issue relevant to the study of racial disparities in legal processing is the existence of qualitative differences in the nature of homicides. For example, if homicides with white victims tend to be premeditated, done by strangers, or committed in the course of another felony, whereas homicides with black victims are found more often to be crimes of passion among friends or relatives, then any demonstrated relationship between race of the victim and the imposition of the death penalty could be explained by differential circumstances of the crimes. Such a qualitative distinction among homicides, based on the victim-offender relationship, has recently been elaborated (Parker and Smith, 1979; Smith and Parker, 1980). "Primary homicides" involve family, friends, or acquaintances and are usually acts of passion; "nonprimary homicides" are most often instrumental, usually occur in the course of another felony, and typi-

cally involve strangers. While intent can be a factor in both types of offenses, this distinction approximates the FBI's legally based classification of "felony" and "non-felony" homicide (FBI, 1974:10; Smith and Parker, 1980). Parker and Smith (1979) have found that variations between states in primary homicide rates are related to poverty and the percentage of the population aged 20–34, while nonprimary rates are related only to the percentage of the population living in urban areas. Because rates of each are associated with a different set of predictors, they argue that the failure to classify homicides in this manner results in incorrect assessment of any correlates of homicide rates.

A third issue of concern is that studies of processing disparities frequently include only a portion of the decisions by the criminal justice system in which racial biases may be found. Most studies have focused their attention on only one decision point, and samples drawn at the later stages of the criminal justice process may be more homogeneous than samples at the entry point (LaFree, 1980; Thomson and Zingraff, 1981). If, for example, white defendants are more likely than blacks to be indicted for a lesser included offense given similar circumstances, such as second versus first degree murder, studies of sentencing severity that do not consider racial differences in indictments can not ask if earlier biases are corrected, and might require highly detailed information about the offender, victim, and circumstances of the crime before any possible sentencing biases are detected (Baldus et al., 1980).

In the present study, data on homicide indictments allow us to assess if race of the defendant, race of the victim, and/or the combination of defendant's and victim's race exert any significant predictive influence on the decision to indict homicide defendants for first versus second or third degree murder and the subsequent decision to impose a death sentence. Associations between these variables will be examined among two groups: among those indicted for any homicide and among only those indicted for first degree murder. In addition, by adopting the distinction between primary and nonprimary homicides, it will be possible to determine if the imposition of the death penalty varies with the nature of the victim-defendant relationship, and if any racial correlates of the death penalty can be explained by the primary/nonprimary distinction.

METHOD

Data were collected in 20 of Florida's 67 counties on all indictments for Murder I, II, and III that occurred in 1976 and 1977. Only defendants found guilty of Murder I, a premeditated homicide or homicide committed in the perpetration of certain felonies, are punishable by death in Florida. Murder II is generally defined as a nonpremeditated homicide resulting from disregard of human life; Murder III is a homicide resulting from certain felonious behavior not designed to effect death (Florida Penal Code, 1979:Chapter 782.04). The counties were selected with the probability of inclusion of each county in the sample proportional to its population size.[1] This procedure led to inclusion of sixteen of Florida's twenty most populous counties—including the eleven largest counties in the state—and four smaller counties. Law students and lawyers visited the courthouse in each county and gathered the case data. To obtain a master list of all homicide indictments, the field investigators reviewed the entire criminal docket in each county for 1976, 1977, and early 1978. Court files for each identified homicide case were then examined, and the researchers completed a standardized information sheet for each. In cases where some of the needed information was missing, the defendant's attorney was contacted and asked to provide relevant details. Local newspapers were also perused for the dates immediately following the offense and trial to obtain any missing data.

Using this procedure, 788 homicide indictments were identified. Because of the salience of the variables measuring race, three cases with an Oriental victim were

[1] These counties were sampled by Professor Hans Zeisel of the University of Chicago.

eliminated. In addition, four cases in which the defendant died before sentencing and six cases in which the defendant was never arrested or jumped bond were also eliminated, reducing the final sample to 775 cases.

For the present analysis, 138 additional cases were deleted because of incomplete information about defendant's race, victim's race, and the relationship between victim and defendant. This reduced the final sample to 637 cases. Most of the deleted cases (n=110) were removed because of an absence of information about the victim-defendant relationship; none was removed in which the death penalty was imposed. A comparison of the deleted cases with those retained reveals that the court records tended to be more complete in cases involving first degree murder indictments or convictions. Whereas 58% of the retained cases involved a first degree murder indictment, only 34% of the deleted cases did so. Similarly, cases with black defendants or victims were slightly more likely to be deleted. However, among the groups indicted for or found guilty of first degree murder in the original sample of 775 cases, neither race of the defendant nor race of the victim was significantly associated with the probability of deletion. Thus, the final sample of 637 cases is representative of the original 775 cases on the variables examined in the analysis.

The variables measuring defendant's and victim's race were dichotomized into white and black categories. Following Parker and Smith (1979), homicides were dichotomized into primary and nonprimary categories on the basis of the victim-defendant relationship. Those classified as primary homicides included those occurring between relatives, friends, ex-lovers, and other primary group members; nonprimary homicides included homicides involving strangers, those who knew each other vaguely but had no friendship or family relationship, and those whose only relationship was through a codefendant (i.e., defendant and codefendant are accused of killing a relative or friend of codefendant). Imposition of the death penalty was measured by coding yes or no, while the severity of indictment was divided into first degree versus second or third degree indictment.

RESULTS

The analysis was designed to address several issues. First, among those indicted for any homicide, we are interested in who is placed in jeopardy of receiving the death penalty by being indicted for first degree murder. Second, for those indicted for any homicide, the analysis will identify those groups having the highest probability of receiving the death penalty. Third, this same analysis will be repeated using only those cases involving a first degree murder indictment. Three possibilities will be assessed in each step. Following Chambliss and Seidman (1971), it is possible that the severity of indictments and sentences are related to the race of the defendant, with a greater proportion of first degree indictments and death sentences imposed on black defendants. Second, race of the victim might be a significant correlate of the severity of charges and sanctions (Bowers and Pierce, 1980; Garfinkle, 1949). Finally, as suggested by Wolfgang and Riedel (1973), the severity of the two legal decisions will be examined to see if they are associated with the joint combination of defendant's and victim's race. This will indicate whether effects of race of the defendant on the severity of indictments or sentences vary between white and black victims, or equivalently whether effects of race of the victim vary between white and black defendants.

Table 1 shows that the probability of being indicted for first degree murder and the probability of receiving the death penalty, once indicted for any homicide, varies with race of the defendant, race of the victim, and the relationship between the defendant and victim. Of the 637 cases, 554 (87%) are intraracial. Looking simply at two-way relationships, there is no significant association between race of the defendant and imposition of the death penalty, as 17/335 (5.1%) of the indicted blacks and 22/302 (7.3%) of the indicted whites were eventually sentenced to death (χ^2 = 1.35). Among black victims, 6/286 were death penalty cases (2.1%), whereas 33/351 (9.4%) of the cases with a white

Table 1. Relationship and Racial Characteristics of Victims and Defendants for All Homicide Indictments

	Number of Cases	First Degree Indictments	Probability of First Degree Indictment	Sentenced to Death	Probability of Death Penalty (all cases)	Probability of Death Penalty (first degree indictments)
Nonprimary						
White victim						
Black defendant	63	58	.921	11	.175	.190
White defendant	151	124	.821	19	.126	.153
Black victim						
Black defendant	103	56	.544	6	.058	.107
White defendant	9	4	.444	0	.000	.000
Primary						
White victim						
Black defendant	3	1	.333	0	.000	.000
White defendant	134	73	.545	3	.022	.041
Black victim						
Black defendant	166	51	.307	0	.000	.000
White defendant	8	4	.500	0	.000	.000
N	637	371	.582	39	.061	.105
Nonprimary						
Black defendant	166	114	.687	17	.102	.149
White defendant	160	128	.800	19	.119	.148
White victim	214	182	.850	30	.140	.165
Black victim	112	60	.536	6	.054	.100

victim resulted in a death sentence ($\chi^2 = 14.63$; $p < .001$). Thus, among all homicide indictments, cases with white victims are more likely to result in a death sentence than are cases with black victims. Furthermore, a greater proportion of black than white victims are involved in primary homicides, as 174/286 cases (60.8%) with black victims involved friends or relatives, as opposed to 137 of the 351 cases (39%) with white victims ($\chi^2 = 30.0$; $p < .001$). Nonetheless, the relationship between race of the victim and death penalty remains when only nonprimary homicides are examined: among this group 5.4% of the cases with a black victim resulted in a death penalty, compared to 14% of the cases with a white victim ($\chi^2 = 5.61$; $p < .05$).

It is also apparent that the death penalty is rarely imposed in cases where the victim and defendant are friends or relatives. Of the primary homicides, only 1% resulted in a death penalty. Because the death penalty is rarely imposed in cases where the victim and defendant had a prior relationship, questions surrounding the imposition of the death penalty are almost solely relevant to cases of nonprimary homicide. Therefore, to avoid obscuring possible correlates and because examining the death penalty among primary homicides produces many zero cells, all further analyses will examine only the 326 cases of nonprimary homicide.

Indictments for First Degree Murder

Despite the disproportionate number of black victims among primary homicide cases and the lower probability of primary homicide cases to result in a first degree murder indictment, there is still a relationship between race of the victim and the decision to indict for first degree murder among nonprimary homicide cases. As presented in Table 1, 53.6% of the nonprimary homicides with a black victim resulted in a first degree murder indictment, whereas 85% of the nonprimary homicides with a white victim involved a first degree indictment. White defendants in non-primary cases are slightly more likely than black defendants to be indicted for first degree murder (80 vs. 68.7%). However, in both the group with black victims and the group with white victims, the probability of a black defendant being indicted for first degree murder is 10% higher than for a white defendant (92.1 vs. 82.1% among white victims and 54.4 vs. 44.4%

among black victims). This reversal is attributable to two associations: between race of the victim and race of the defendant and between race of the victim and indictment for first degree murder. That is, the overall higher probability for white defendants to be indicted for first degree murder is a function of the tendency of whites to kill other whites. Similarly, a black accused of killing a white has a 37.7% higher probability of being indicted for first degree murder than a black suspected of killing a black (92.1 minus 54.4%).

The analysis of the data on nonprimary homicides presented in Table 1 can be simplified by use of log-linear techniques (Burke and Turk, 1975; Fienberg, 1980; Swafford, 1980). This procedure, which assumes multiplicative effects (linear effects on a log scale) of variables on cell probabilities, compares expected or hypothesized frequencies representing a particular model to those frequencies actually observed in the data in terms of a likelihood-ratio χ^2 test. Its goal is to build a model that fits the data so that cell frequencies can be reproduced (within limits of sampling error). Interaction terms between variables can be added or deleted until the most parsimonious model fitting the data is discovered. The three variables in Table 1 are referred to as D (defendant's race), V (victim's race), and F (first degree murder indictment). Thus, the F,V,D Model hypothesizes that each variable is independent of the other two; the FV,VD Model postulates that a first degree indictment is correlated with victim's race and victim's race is correlated with race of the defendant, but that first degree indictment and defendant's race are independent when controlling for victim's race (see Swafford, 1980 for further elaboration of model notation).

Table 2 presents 8 models that were fit to the three variables in Table 1.[2] The first model, as described above, postulates that the three variables are mutually independent. This model is rejected because the probability of obtaining a $\chi^2 \geq 167.32$ if this model were indeed correct is $< .0001$. The eighth model, postulating associations between each pair of variables (FV,FD, VD) does describe the observed data nicely, with a χ^2 of .34 (p=.5616). The only possible effect omitted from this model is a three-factor interaction among the variables; the lack of significance of this omission indicates that the relationship between defendant's race and magnitude of indictment does not significantly vary with race of the victim. If the FD term is eliminated from Model 8, as done in Model 6, χ^2 increases 3.42 and 1 degree of freedom, which does not attain statistical significance at the .05 level. If Model 6 were indeed the true model, the probability of obtaining a $\chi^2 \geq 3.76$ is .15. Thus, Model 6 appears to be the most parsimonious model having an adequate fit. This indicates that the race of victims and defendants is intercorrelated (VD), that nonprimary homicides with white victims are more likely than nonprimary homicides with black victims to result in an indictment for first degree murder (FV), and that, controlling for victim's race, severity of indictment and defendant's race are independent. In short, there is a tendency to indict defendants accused of killing whites for first degree murder and to indict defendants accused of killing blacks for a less severe homicide charge. However, the defendant's race does not have any statistically significant effects.

Table 2. Models Fit to the Three-Variable Contingency Table Examining All Indictments and First Degree Murder Indictment as Outcome Variable

Model	Fitted Marginals	Likelihood Ratio χ^2	D.F.	Probability
M1	F,V,D	167.32	4	.0
M2	FV,D	130.84	3	.0
M3	FD,V	161.93	3	.0
M4	VD,F	40.24	3	.0
M5	FV,FD	125.45	2	.0
M6	FV,VD	3.76	2	.1525
M7	FD,VD	34.85	2	.0
M8	FV,FD,VD	.34	1	.5616

NOTE: F = First Degree Murder Indictment; V = Victim's Race; D = Defendant's Race.

[2] Before undertaking the log-linear analysis, .5 was added to each cell in the contingency table, which slightly reduces the strength of the partial associations.

The Imposition of the Death Penalty

Returning to Table 1, both race of the defendant and race of the victim appear to be related to the imposition of the death penalty among nonprimary homicides. Overall, white defendants are slightly more likely to be sentenced to death (11.9 vs. 10.2%; not significant), although among both cases with a white victim and cases with a black victim the probability of receiving the death penalty is approximately 5% higher for black defendants (17.5 vs. 12.6% for white victims and 5.8 vs. 0% for black victims). Both white and black defendants have a 12% higher probability of receiving the death penalty if they are accused of killing a white instead of a black (12.6 vs. 0% for white defendants and 17.5 vs. 5.8% for black defendants). Again, because of different cell sizes and the intercorrelation between victim's and defendant's race, log-linear analysis is used to identify the most salient relationships.

Table 3 presents 8 models that were fit to the contingency table, where Y = death penalty, V = victim's race, and D = defendant's race. All cases of nonprimary homicide are included. The eighth model, postulating associations between each pair of variables, describes the data very well, with a χ^2 of .01 (p = .9194).[3] Because the three-factor interaction is omitted from this model, it can be concluded that any relationship between the death penalty and race of the victim does not significantly vary with race of the defendant. This model, however, can be further simplified. Model 6 eliminates the YD term from Model 8, hypothesizing that the death penalty, given race of the victim, is conditionally independent of race of the defendant. Because χ^2 increases by only .95 and one degree of freedom, this YD term can be safely dropped. χ^2 increases by an additional 5.41 and one degree of freedom when the YV term is also dropped, so Model 4 is significantly poorer than Model 6. The other models also give inadequate fits, thus indicating that the most useful model in describing the data is Model 6. This model indicates that there are two significant partial relationships, YV and VD, neither of which significantly varies by levels of the third variable. Thus, the imposition of the death penalty is related to race of the victim controlling for race of the defendant, and an intercorrelation exists between race of the defendant and race of the victim. Any possible effects of race of the defendant on the imposition of the death penalty (controlling for victim's race) or the possibility that the effects of race of the victim vary with race of the defendant do not add any significant descriptive power to the model.

Table 3. Models Fit to the Three-Variable Contingency Table Examining All Indictments and Death Penalty as Outcome Variable

Model	Fitted Marginals	Likelihood Ratio χ^2	D.F.	Probability
M1	Y,V,D	133.45	4	.0
M2	YV,D	128.04	3	.0
M3	YD,V	133.24	3	.0
M4	VD,Y	6.37	3	.0947
M5	YV,YD	127.82	2	.0
M6	YV,VD	.96	2	.6186
M7	YD,VD	6.16	2	.0460
M8	YV,YD,VD	.01	1	.9194

NOTE: Y = Death Penalty; V = Victim's Race; D = Defendant's Race.

This same analysis was repeated for the 242 cases which involved nonprimary homicides and first degree murder indictments.[4] Because of the reduction in sample size, stronger relationships are necessary in this step to attain statistical significance. As displayed in Table 4, the model postulating all three partial associ-

[3] This model, with all two-factor effects present, can be used to compare the probabilities of receiving the death penalty while controlling for one variable. Using the procedure described by Fienberg (1980:16–9), and controlling for race of the victim, the estimated odds of a white receiving a death sentence are .68 the odds of a black. Controlling for race of the defendant, the estimated odds of receiving the death penalty are 3.22 times higher if the victim is white than if the victim is black.

[4] This step would not be necessary if a 4-dimensional table with F, Y, V, and D was originally analyzed. This is not done because of the impossibility of receiving the death penalty if not indicted for first degree murder, thereby producing an incomplete table with numerous zero cells.

ations again fits well, indicating no significant three-factor interaction. The YD term can be deleted from this model with no significant loss of fit, as done in Model 6, thus resembling the model selected when all nonprimary homicide indictments were examined. However, this model can be further simplified by deleting the YV term, as done in Model 4. Because χ^2 increases by only 1.16 and one degree of freedom over that observed in Model 6, which is not significant, Model 4 appears to be the best fitting model. This indicates that among those indicted for first degree murder, the only significant relationship is between race of the defendant and race of the victim, and neither of these two variables nor their interaction is a significant correlate of the imposition of a death sentence.[5]

Among this group of those indicted for first degree murder, the lack of statistical significance of the tendency to impose the death penalty on those with white victims more frequently than on those with black victims must be interpreted with caution. As shown in Table 1, 16.5% of the 182 cases with a white victim in this group and 10% of the 60 cases with a black victim resulted in a death sentence. These trends continue the patterns observed when the type of indictment was treated as the dependent variable, and are consistent with patterns observed in other death-penalty studies (Bowers and Pierce, 1980; Garfinkle, 1949). A 95% confidence interval placed on this 6.5% observed difference indicates that actual differences in the population of all those indicted for first degree murder fall between −2.8% to 15.8% (Agresti and Agresti, 1979:172). This interval is wide because of relatively small sample sizes, and its inclusion of zero prohibits rejection of the possibility that no racial differences actually exist. Thus, insufficient evidence of an association does not necessarily mean that there is no association between the variables in the population. Evidence from other studies, taken together with the range of this confidence interval, suggests that an association between race of the victim and a sentence of death, given an indictment for first degree murder, could be found if the sample size was increased.

[5] A similar model is found when only those cases (n = 94) involving a conviction for first degree murder are examined.

Table. 4. Models Fit to the Three-Variable Contingency Table Examining First Degree Murder Indictments Only and Death Penalty as Outcome Variable

Model	Fitted Marginals	Likelihood Ratio χ^2	D.F.	Probability
M1	Y,V,D	76.19	4	.0
M2	YV,D	75.03	3	.0
M3	YD,V	76.19	3	.0
M4	VD,Y	1.59	3	.6607
M5	YV,YD	75.03	2	.0
M6	YV,VD	.43	2	.8049
M7	YD,VD	1.59	2	.4508
M8	YV,YD,VD	.01	1	.9402

NOTE: Y = Death Penalty; V = Victim's Race; D = Defendant's Race.

SUMMARY AND CONCLUSIONS

Among primary homicides in twenty Florida counties in 1976 and 1977, the data indicate that those accused of murdering whites have a significantly higher probability of being placed in jeopardy of receiving the death penalty by being indicted for first degree murder than those accused of murdering blacks. Similarly, those accused of murdering whites are significantly more likely to eventually receive the death penalty than those accused of murdering blacks. However, once indicted for first degree murder, neither race of the defendant, race of the victim, nor the interaction between these two variables are significantly related to the probability of receiving the death penalty. Thus, the higher probability of being sentenced to death for those accused of killing whites among the group indicted for any nonprimary homicide can be seen primarily as a function of the higher probability of those accused of killing whites being indicted for first degree murder. Once indicted for first degree murder, the evidence from this sample is not strong enough to conclude that those accused of killing whites are significantly more likely to be sentenced to death. It can therefore be concluded that given a homicide, deci-

sions made by the prosecutor and grand jury place those accused of murdering whites in greater jeopardy of receiving the death penalty through a higher probability of being indicted for first degree murder.

Contrary to the theoretical arguments of Black (1976), Chambliss and Seidman (1971), and Quinney (1979), and consistent with the findings of Hagen (1974), there are no strong differences in this sample between white and black defendants in either the probability of being indicted for first degree murder or in the overall probability of being sentenced to death. However, when broken down by race of the victim, black defendants are more likely to be indicted for first degree murder and sentenced to death among both categories of victim's race. Nevertheless, there is not enough evidence to conclude that the interaction between victim's race and defendant's race is a significant correlate of receiving a death sentence. Although these trends might have substantive importance, the only statistically significant correlate of the death penalty among the 326 cases is race of the victim.

The major finding that remains unexplained is why nonprimary homicides with white victims are more likely to result in an indictment for first degree murder than are nonprimary homicides with black victims. This trend has the effect of disproportionately channelling cases with white victims toward the death penalty. Despite this apparent discrepancy, such prosecutorial discretion tends to be less visible and more immune to judicial review than are racial disparities in sentencing severity.

It is clear that studies examining only sentencing severity or only defendant characteristics (e.g., Chiricos and Waldo, 1975) risk missing significant racial biases in legal processing. The strength of the racial disparities observed in this study will fluctuate as other potentially relevant variables are introduced, such as the sex, age, and criminal histories of the victim and defendant or region of the state. Some research has suggested that racial factors may operate through social class (Green, 1970; Farrell and Swigert, 1978), since the higher representation of blacks in the lower class makes them disproportionately susceptible to any class-based inequalities. Finally, the division of nonprimary homicides into additional subcategories will allow more precise specification of the relationship between racial characteristics and legal processing (Baldus et al., 1980).

In conclusion, relative equality in the imposition of the death penalty appears mythical as long as prosecutors are more likely to obtain first degree murder indictments for those accused of murdering white strangers than for those accused of murdering black strangers. Racial differences in the processing of those indicted for nonprimary homicides in Florida appears to place a lower value on the lives of blacks than on the lives of whites.

REFERENCES

Agresti, Alan and Barbara Finlay Agresti
1979 Statistical Methods for the Social Sciences. San Francisco: Dellen.

Baldus, David C., Charles A. Pulaski, George Woodworth, and Frederick D. Kyle
1980 "Identifying comparatively excessive sentences of death: a quantitative approach." Stanford Law Review 33:1–74.

Bernstein, Ilene Nagel, William R. Kelly, and Patrica A. Doyle
1977 "Societal reaction to deviants: the case of criminal defendants." American Sociological Review 42:743–55.

Black, Donald
1976 The Behavior of Law. New York: Academic.

Bowers, William J.
1974 Executions in America. Lexington, MA: Lexington Books.

Bowers, William J. and Glenn L. Pierce
1980 "Arbitrariness and discrimination in post-Furman capital cases." Crime and Delinquency 26:563–635.

Burke, Peter J. and Austin T. Turk
1975 "Factors affecting postarrest dispositions: a model for analysis." Social Problems 22:313–32.

Chambliss, William J. and Robert Seidman
1971 Law, Order, and Power. Reading, MA: Addison-Wesley.

Chiricos, Theodore G. and Gordon P. Waldo
1975 "Socioeconomic status and criminal sentencing: an empirical assessment of a conflict proposition." American Sociological Review 40:753–72.

Farrell, Ronald A. and Victoria Lynn Swigert
1978 "Legal disposition of inter-group and intra-group homicides." Sociological Quarterly 19:565–76.

Federal Bureau of Investigation
1974 Uniform Crime Reports, 1973. Washington, D.C.: U.S. Government Printing Office.
Fienberg, Stephen E.
1980 The Analysis of Cross-Classified Categorical Data, 2nd Ed. Cambridge, MA: MIT Press.
Garfinkle, Harold
1949 "Research note on inter- and intra-racial homicides." Social Forces 27:369–81.
Green, Edward
1970 "Race, social status, and criminal arrest." American Sociological Review 35:476–90.
Hagen, John
1974 "Extra-legal attributes and criminal sentencing: an assessment of a sociological viewpoint." Law and Society Review 8:357–84.
Hindelang, Michael J.
1969 "Equality under the law." The Journal of Criminal Law, Criminology, and Police Science 60:306–13.
1978 "Race and involvement in common law personal crimes." American Sociological Review 43:93–109.
LaFree, Gary D.
1980 "The effects of sexual stratification by race on official reactions to rape." American Sociological Review 45:842–54.
Myers, Martha A.
1979 "Offended parties and official reactions: victims and the sentencing of criminal defendants." Sociological Quarterly 20:529–40.
Parker, Robert Nash and M. Dwayne Smith
1979 "Deterrence, poverty, and type of homicide." American Journal of Sociology 85:614–24.
Quinney, Richard
1979 Criminology, 2nd Ed. Boston: Little, Brown, and Co.
Riedel, Marc
1976 "Discrimination in the imposition of the death penalty: a comparison of the characteristics of offenders sentenced pre-Furman and post-Furman." Temple Law Quarterly 49:261–86.
Sellin, Thorsten
1980 The Penalty of Death. Beverly Hills: Sage.
Smith, M. Dwayne and Robert Nash Parker
1980 "Type of homicide and variation in regional rates." Social Forces 59:136–47.
Swafford, Michael
1980 "Three parametric techniques for contingency table analysis: a nontechnical commentary." American Sociological Review 45:664–90.
Swigert, Victoria Lynn and Ronald A. Farrell
1977 "Normal homicides and the law." American Sociological Review 42:16–32.
Thomson, Randall J. and Matthew T. Zingraff
1981 "Detecting sentencing disparity: some problems and evidence." American Journal of Sociology 86:869–80.
U.S. Department of Justice
1980 Capital Punishment, 1979. Washington, D.C.: U.S. Government Printing Office.
Unnever, James D., Charles E. Frazier, and John C. Henretta
1980 "Race differences in criminal sentencing." Sociological Quarterly 21:197–206.
Wolfgang, Marvin E. and Marc Riedel
1973 "Race, judicial discretion, and the death penalty." Annals of the American Academy of Political and Social Science 407:119–33.

COMPLEXITIES OF THE RANDOMIZED RESPONSE SOLUTION*

(COMMENT ON TRACY AND FOX, *ASR*, APRIL 1981)

JUDITH DROITCOUR MILLER
The George Washington University

As most sociologists and criminologists are aware, a problem for surveys of deviant behavior has been that while "innocent" respondents provide truthful answers to sensitive questions, many deviants deny the facts of their experience or behavior. The randomized response technique first suggested by Warner in 1965 seems appealing because it is designed to encourage truthful responses on the part of deviants. However, as highlighted by the results Tracy and Fox have presented in a recent article (1981), little is gained if the technique unwittingly encourages the "innocent" to falsify their answers. This risk pertains whenever nondeviant respondents fear that a randomly assigned response could be interpreted (erroneously) as meaning that they have committed a deviant act. Unless such fears are neutralized, a randomized response estimate of deviant behavior prevalence or frequency can hardly be deemed more valid than a direct-question estimate. "Counterbalancing" the "risks of suspicion" which innocent respondents incur and the reduced jeopardy afforded to deviant respondents is hardly a sufficient solution. This would be obvious in surveys of general population groups, since the large majority of such respondents typically have not engaged in the deviant behavior which is the focus of the study. While it is heartening to see the sociological recognition of response-validity problems (and of new ideas for their solution) implied by inclusion of the Tracy-Fox article in this journal, some of the results deserve clarification and further discussion, while others are marked by problems of which readers should be made aware.

Briefly, the Tracy-Fox validity study consisted of interviews with respondents known to have been arrested one or more times. A portion of these interviews were conducted in the conventional manner, while the remainder featured the randomized response approach for the sensitive question concerning the number of times respondent had been arrested. Respondents who drew a plain red ball were to give the number of times they had been arrested, while respondents who drew a numbered white ball were to say the number on the ball. Numbers on white balls ranged from 0 to 8. Results showed that for those who had been arrested on multiple occasions, the randomized response estimate was, as expected, considerably higher and closer to the mean of officially recorded arrests than was the direct-question mean. But for respondents who were arrested on a single occasion, the randomized response condition produced an estimate of only .27 arrests per person, in comparison to the mean of .76 which was obtained with direct questioning. The two very different effects that randomized response had on the bias of estimates for the two arrestee groups is reported in Table 2 of the Tracy-Fox article. However, the potential import of these contrasting effects—and the fact that they dwarf the "main effect" observed for the total sample—becomes abundantly clear only when the results are presented graphically, as shown in Figure 1 below.

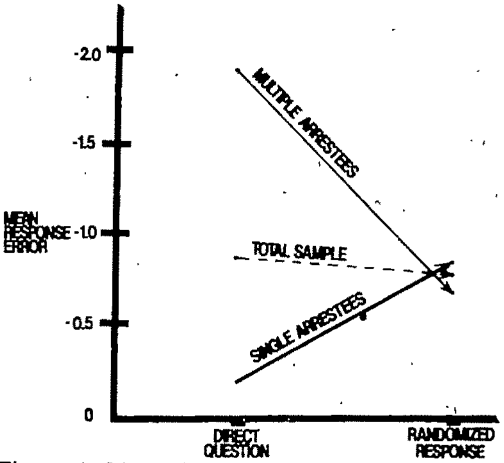


Figure 1. Mean Response Error by Interview Method, for Two Arrestee Groups

As indicated by the crossed lines in this graph, the randomized response technique succeeded in reducing response bias for multiple arrestees, but served to increase response bias for single arrestees (the relative "innocents"). Thus, the direction as well as the size of the overall (total sample) effect of randomized response is dependent upon the relative sizes of the two arrestee groups. As Tracy and Fox proceed with a total-sample comparison of the randomized response and direct-question results, they fail to stress this point, although it is clearly suggested by the various

* Direct all correspondence to Judith Droitcour Miller, Social Research Group, The George Washington University, 2401 Virginia Ave., N.W., Washington, DC 20037.

results they present based on assignment of different relative weights to the single and multiple arrestee groups. Unfortunately, in each case (including the use of "unweighted" figures or natural sample Ns), multiple arrestees are over-represented relative to their prevalence in the population from which the sample was drawn. The implication seems to be that were these data weighted to represent the actual arrestee population, the reduction in response bias achieved by randomized response would be minimal or perhaps even nonexistent. Indeed, in a general population survey, where the majority of persons have never been arrested, the expectation would be that the Tracy-Fox procedures of randomized response would actually increase the overall level of response bias. Since the outcome of employing the Tracy-Fox procedures is directly dependent on population composition, the generalizability implied in several of their conclusions seems to be inappropriate. This is particularly true for the table comparing levels of mean-square error for direct-question and randomized response estimates at various sample-size levels.

The interesting questions posed by these data are: How can we explain the contrasting results obtained for single and multiple arrestees? And, how might it be possible to avoid the problems of response validity Tracy and Fox encountered for single arrestees?

It seems reasonable to assume that under the randomized response condition, respondents who drew plain red balls (and who were therefore required to voice the actual number of times they had been arrested) would be at least as willing to provide a truthful answer as were those who participated in a conventional, direct-question interview. Under this assumption, the estimates obtained for single arrestees can only be explained by false responses on the part of those who drew numbered white balls. The logical explanation—and incidentally, one which is consistent with Goffman's (1959) theory of the presentation of self—would be that single arrestees who drew high-numbered white balls were reluctant to misrepresent their behavior or experience in such a way as to cast a negative light upon themselves. Thus, although detailed distributions were not provided by Tracy and Fox, it seems likely that some of the single arrestees who drew high numbers disobeyed the response rules and either provided a true answer to the sensitive question (i.e., said "one" for one arrest) or took advantage of this opportunity to rule out suspicion (by saying "zero"). As Tracy and Fox briefly noted,

> Stratifying the comparisons according to frequency of official arrests . . . highlights the substantial effects of [differential] respondent hazards, i.e., jeopardy and risk of suspicion. . . . Corollary to protecting respondents from having to reveal sensitive information is permitting a respondent who is relatively innocent to appear as such. In other words, compelling a respondent to give an answer that provokes undue suspicion should be avoided. [1981:194]

Part of the response validity problem Tracy and Fox encountered in their application of the randomized response technique can be explained by the fact that the sensitive question constituted the only meaningful interpretation of the response which was shared by both interviewer and respondent. An alternative model of randomized response, briefly described in the Tracy-Fox article, allows the investigator to avoid—or at least to mitigate—this sort of "respondent jeopardy." In the unrelated-questions model of randomized response (which has been developed primarily by Greenberg, see e.g., Greenberg, Abul-Ela, Simmons, and Horvitz, 1969; Greenberg, Kuebler, Abernathy, and Horvitz, 1971), respondents are presented with two substantive questions: the sensitive question and an innocuous question. The secret random draw merely determines which of the two substantive questions the respondent will answer. This version of randomized response seems to be superior in that it provides both respondent and interviewer with a meaningful alternative interpretation of whatever answer the respondent may voice. Studies of highly stigmatized behaviors may require more powerful neutralization techniques, such as using a socially positive innocuous item. For example, Zdep and Rhodes (1976–77) found it necessary to pair a sensitive child-abuse item with socially positive items such as PTA attendance and church/synagogue attendance.

In hindsight it might appear that Tracy and Fox should have paired their sensitive item with a question such as "How many times have you stopped to help a stranger whose car had broken down?" But to do so would have been rather costly in terms of the variance of the randomized response estimate. Because the numbers on the white balls were determined by Tracy and Fox, they were able to set the mean of this distribution equal to their expectation for the study-group mean on the sensitive question; and of course, they knew the population mean of the white-ball distribution. In contrast, if they had chosen a socially positive, innocuous item, it is unlikely that the population mean would be known; thus, a sample mean would have to serve as an estimate of the innocuous item and such an estimate would contribute additional variance to the randomized response estimate. Moreover, it is

unlikely that even a sample mean for the innocuous item would be available in advance (at least not for a sample matched to the study group), so it would have been necessary to devote a portion of the sample to estimating the innocuous mean. Lastly, if the innocuous-item mean does not turn out to be close to the sensitive-item mean, the variance of the randomized response estimate is increased considerably (see Pollack and Bek, 1976). As these considerations indicate, the validity-reliability tradeoffs facing investigators considering employing the randomized response technique are more complex than is suggested in the Tracy-Fox article.

The "fatal flaw" of the randomized response approach seems to be that all too often the design decision which would favor response validity carries an unacceptably high variance cost (see Miller and Cisin, 1980). For this reason, the more fruitful paths for future development may lie in the direction of alternative methods of indirect, survey-based estimation. One such approach is "aggregated response" (see Boruch and Cecil, 1979, citing Raghavarao and Federer, 1973). Another new approach is the nominative technique, a variation on the multiplicity techniques developed by Sirken (see Sirken, 1975; Fishburne, 1980; Fishburne, Abelson, and Cisin, 1980; Cisin, 1980). Yet another new technique, the "item-count/paired lists" approach, is currently being developed at The George Washington University under a grant from the National Institute on Drug Abuse.

REFERENCES

Boruch, Robert F. and Joseph S. Cecil
1979 Methods for Assuring Privacy and Confidentiality of Social Research Data. Philadelphia: University of Pennsylvania Press.

Cisin, Ira H.
1980 "The variance of prevalence estimates." Pp. 65–73 in Joan D. Rittenhouse (ed.) Developmental Papers: Attempts to Improve the Measurement of Heroin Use in the National Survey—An Administrative Report. Rockville, Md.: National Institute on Drug Abuse.

Fishburne, Patricia M.
1980 Survey techniques for studying threatening topics: A case study on the use of heroin. Ph.D. Dissertation, Department of Sociology, New York University.

Fishburne, Patricia M., Herbert I. Abelson, and Ira H. Cisin
1980 National Survey on Drug Abuse: Main Findings, 1979. Washington, D.C.: Superintendent of Documents, U.S. Government Printing Office. DHHS Publication No. (ADM) 80–976.

Goffman, Erving
1959 The Presentation of Self in Everyday Life. Garden City, N.Y.: Doubleday-Anchor Books.

Greenberg, Bernard G., Abdel-Latif A. Abul-Ela, Walt R. Simmons, and Daniel G. Horvitz
1969 "The unrelated question randomized response model theoretical framework." Journal of the American Statistical Association 64:520–39.

Greenberg, Bernard G., Roy R. Kuebler, Jr., James R. Abernathy, and Daniel G. Horvitz
1971 "Application of the randomized response technique in obtaining quantitative data." Journal of the American Statistical Association 66:243–50.

Miller, Judith D. and Ira H. Cisin
1980 "Measuring the hidden prevalence of heroin use." Pp. 74–170 in Joan Dunne Rittenhouse (ed.) Developmental Papers: Attempts to Improve the Measurement of Heroin Use in the National Survey—An Administrative Report. Rockville, Md.: National Institute on Drug Abuse.

Pollack, K. H., and Yuksel Bek
1976 "A comparison of three randomized response models for quantitative data." Journal of the American Statistical Association 71:884–6.

Sirken, Monroe G.
1975 "Network surveys of rare and sensitive conditions." Pp. 31–2 in Advances in Health Survey Research Methods, National Center on Health Statistics Research Proceedings Series.

Tracy, Paul E. and James Alan Fox
1981 "The validity of randomized response for sensitive measurements." American Sociological Review 46:187–200.

Warner, Stanley L.
1965 "Randomized response: a survey technique for eliminating evasive answer bias." Journal of the American Statistical Association 60:63–9.

Zdep, S. M. and Isabelle N. Rhodes
1976–77 "Making the randomized response technique work." Public Opinion Quarterly 40:531–7.

REAFFIRMING THE VIABILITY OF THE RANDOMIZED RESPONSE APPROACH*

(REPLY TO MILLER)

JAMES ALAN FOX
Northeastern University

PAUL E. TRACY
University of Pennsylvania

The randomized response approach—like many methodological advances which present an alternative to, rather than an extension of,

* Direct communication to James Alan Fox, College of Criminal Justice, Northeastern University, Boston, MA 02115 or Paul E. Tracy, Center for Studies in Criminology and Criminal Law, University of Pennsylvania, Philadelphia, PA 19104.

traditional methods—was confronted with considerable resistance in the social sciences after its introduction by Warner (1965).[1] While naive skeptics questioned how purposely imposing error into an interview design could be at all beneficial, wiser observers awaited evidence that the introduction of (random) error could be balanced by a sufficiently large reduction in response bias. It is of little surprise, therefore, that most field applications of the technique have been comparative (contrasting randomized response against traditional survey approaches) rather than substantive.[2]

The field-validation efforts that have been accomplished (see, for example, Folsom, 1974; Locander et al., 1974; Bradburn and Sudman, 1979; Tracy and Fox, 1981; and Cahalan and Ekstrand, 1980) appear to indicate that, when appropriately applied, randomized response can enhance our ability to survey sensitive topics. It is clear from reviewing the validation literature that the performance of the randomized response approach (or, in fact, most any procedure) depends on how well-suited the specific design features are for the population and topic under investigation. That is, the randomized response approach defines a wide range of survey response plans (limited only by the imagination of the investigator) that vary in terms of certain statistical properties (e.g., efficiency) as well as practicality; the design and the execution of a randomized response survey (like any survey) determine the quality of the responses that it generates.

Our validation results (Tracy and Fox, 1981) favored overall the use of randomized response. However, one qualification, which we indicated and which received needed emphasis in Miller's comment on our research findings, concerns the poor performance of our randomized response procedure with the relatively "innocent," i.e., respondents with only one arrest.[3]

Based on the available pool of arrestees (having a mean of arrests approaching three), we had anticipated that after some uneven slippage in locating arrestees the final interviewed sample would average nearer to two in official arrests. The randomized response procedure was designed accordingly: the white balls in the randomizing devices averaged 2.2 (by specification) to provide response protection for the anticipated sample. However, success in locating the arrestees was more uneven than expected, producing an overabundance of one-time arrestees and a mean of arrests (in the randomized response condition) of only 1.45. The three sets of results shown in Table 3 (which all support randomized response to some degree) should illustrate the effect that suitability of randomized response technique to the sample composition can have. That is, had the white ball distribution been more suitable (i.e., having a mean set between one and two), the abundant, "innocent" group would have been granted greater protection and, thus, the results both for this group and overall would have improved. Therefore, although Miller stresses the effect that sample composition has on overall performance of the procedure, we stress the converse: the effect of instrument suitability to sample composition.

Miller also remarks that in a general population where the majority of persons have never been arrested, the results of our procedure would have increased response bias. However, besides the fact that a general population survey of arrests (as opposed to self-reported crime or victimization) would be of little substantive interest, the randomized response procedure (such as the random digits we employed) would in any sensitive survey be adjusted for anticipated levels of prevalence or incidence of the attributes or behaviors under investigation.

After discussing the particular results of our research, Miller recommends a fascinating alternative to the random digits that we used to protect responses: a nonsensitive question that elicits socially desirable responses should destigmatize further a high response by attaching to it a possible desirable meaning rather than

[1] In contrast, methodological innovations that extend traditional approaches (as log-linear models did with crosstabular analysis and as structural equation models did with regression analysis) are more quickly embraced by applied researchers.

[2] In the field of criminology, for example, conventional wisdom has thus far yielded to the point where randomized response is suggested as a validation check on traditional surveys of victims, but not as a primary data collection device (see Penick and Owens, 1976:74).

[3] Actually, the poor results for the one-time arrestees should not be surprising. In attempting to protect those who are asked to provide sensitive responses, the randomized response technique, by definition, risks implication for those who are innocent. (For more discussion of these respondent hazards with suggestions for handling them, see Leysieffer and Warner, 1976; Greenberg et al., 1977; Fox and Tracy, 1980). In the particular response plan we employed, respondents were instructed to select at random a ball from a device containing 25 red balls and 25 white balls numbered zero through eight, and to report either their frequency of arrests or a random digit depending on the outcome of the selection. As a consequence, one-time arrestees would be more uncomfortable than the multiple arrestees with the suspicion created if directed by the randomized response device to report a large random digit.

just a possible neutral meaning (as with the random digits). Although intrigued, we do not see the basis for Miller's unsupported claim that this approach is "superior."

First, as Miller concedes, her proposition is clearly inferior in terms of statistical efficiency. In contrast to random digits which follow a prespecified known distribution, the distribution (mean and variance) of the responses to the socially desirable question would need to be estimated. This task could be accomplished by devoting a portion of the sample to a direct survey of the nonsensitive question (see Moors, 1971) or, more efficiently, by using two nonsensitive questions which alternately in split-halves of a sample are either posed directly (to estimate the distribution of the nonsensitive responses) or paired with the sensitive question in a randomized response scheme (see Folsom et al., 1973; Zdep and Rhodes, 1976–77). With either procedure, sampling error in the estimates of the distributional properties of the nonsensitive question(s) inflates the variance of the parameter estimates for the sensitive question. As Miller notes, the extent of variance inflation increases the more the true means of the sensitive and nonsensitive responses differ. Moreover, the variance of the nonsensitive responses must be sufficiently large to provide response protection along the range of possible sensitive responses; but if too large, the variances of the sensitive response estimates further inflate. It was these difficulties in inventing a meaningful nonsensitive question having the "right" distribution that would not unnecessarily inflate the estimation variance, as well as the increased variance produced simply by having to estimate this distribution, that led us to elect a random digit "nonsensitive question."

More serious than the reliability issues discussed above is a potential bias created by using a nonsensitive question that solicits socially desirable responses. In pursuing an estimate of the distributional properties of the nonsensitive question, respondents would tend to exaggerate in answering a direct inquiry. However, when this same nonsensitive question were paired with the sensitive question in a randomized response inquiry, the tendency to overreport on a nonsensitive question such as "How many times have you stopped to help a stranger whose car had broken down?" would be discouraged by the possible undesirable interpretation of the response attached by the sensitive question. In other words, a direct estimate of the mean of the desirable nonsensitive responses would overstate the mean actually operative under the more threatening randomized response condition. Consequently, an estimate of the mean of the sensitive responses, defined as a function of the inflated estimate of the nonsensitive response mean, would be seriously understated.

In brief, Miller's suggestion is appealing at one level. However, it is not clear that the advantage of using meaningful and desirable alternative questions over meaningless and neutral random digits outweighs the complex estimation problems that arise when using alternative questions whose distributions are unknown. Moreover, because the social undesirability of a sensitive question will often far exceed the social undesirability of a nonsensitive question, the advantage of employing socially desirables may even be small. Again, Miller's assertion that this approach is "superior" is premature.

Miller's comment concludes with a brief mention of alternative designs for protecting respondents in sensitive surveys. Indeed, the developmental literature regarding randomized response models and some offshoots such as linear random transformations of sensitive responses (e.g., Warner, 1971; Himmelfarb and Edgell, 1980) and balanced incomplete block total responses (see Raghavarao and Federer, 1979) is progressing in interesting directions. Unfortunately, very few field tests have appeared comparing alternative response plans or comparing strategies for structuring the features of randomized response plans.[4] We disagree that randomized response necessarily has a "fatal flaw;" simply, more needs to be understood about the reactions of respondents—including their comfort, cooperation, and comprehension—to enable the design of optimal randomized response plans.

REFERENCES

Bradburn, Norman M. and Seymour Sudman
1979 Improving Interview Method and Questionnaire Design. San Francisco: Jossey-Bass.

Cahalan, Margaret and Laurie E. Ekstrand
1980 "Who are the tax cheaters and why do they cheat?" Rockville, MD: Westat, Inc.

Folsom, Ralph E.
1974 "A randomized response validation study: comparison of direct and randomized reporting on DUI arrests." Research Triangle Park, NC: Research Triangle Institute.

Folsom, Ralph E., Bernard G. Greenberg, Daniel G. Horvitz, and James R. Abernathy
1973 "The two alternative questions randomized response model for human surveys." Journal of the American Statistical Association 68:525–30.

[4] Notable exceptions are the field test, albeit small in scale, of several quantitative randomized response models offered by Smith et al. (1974) and Moriarty and Wiseman's (1977) field comparison of several randomization devices.

Fox, James Alan and Paul E. Tracy
1980 "The randomized response approach: applicability to criminal justice research and evaluation." Evaluation Review 4:601–22.

Greenberg, Bernard G., Roy R. Kuebler, James R. Abernathy, and Daniel G. Horvitz.
1977 "Respondent hazards in the unrelated question randomized response model." Journal of Statistical Planning and Inference 1:53–60.

Himmelfarb, Samuel H., and Stephen E. Edgell
1980 "The additive constants model: a randomized response technique for eliminating evasiveness to quantitative response questions." Psychological Bulletin.

Leysieffer, Frederick W. and Stanley L. Warner
1976 "Respondent jeopardy and optimal designs in randomized response models." Journal of the American Statistical Association 71:649–56.

Locander, William B., Seymour Sudman, and Norman M. Bradburn
1974 "An investigation of interview method, threat and response distortion." Journal of the American Statistical Association 71:269–75.

Moors, J. J. A.
1971 "Optimization of the unrelated question randomized response model." Journal of the American Statistical Association 66:627–9.

Moriarty, Mark and Frederick Wiseman
1976 "On the choice of a randomization technique with the randomized response model." Proceedings of the Social Statistics Section of the American Statistical Association:624–6.
1976 Surveying Crime. Washington, D.C.: National Academy of Sciences.

Raghavarao, Damaraju and Walter T. Federer
1979 "Block total response as an alternative to the randomized response method in surveys." Journal of the Royal Statistical Society, Series B 41:40–5.

Smith, Linda L., Walter T. Federer, and Damaraju Raghavarao
1974 "A comparison of three techniques for eliciting truthful answers to sensitive questions. Proceedings of the Social Statistics Section of the American Statistical Association:447–52.

Tracy, Paul E. and James Alan Fox
1981 "The validity of randomized responses for sensitive measurements." American Sociological Review 46:187–200.

Warner, Stanley L.
1965 "Randomized response: a survey technique for eliminating evasive answer bias." Journal of American Statistical Association 60:63–9.
1971 "The linear randomized response model." Journal of the American Statistical Association 66:884–8.

Zdep, S. M. and Isabelle N. Rhodes
1976–77 "Making the randomized response technique work." Public Opinion Quarterly 40:531–7.

ON INTERPRETING COEFFICIENTS IN DIFFERENTIAL EQUATION MODELS

(COMMENT ON NIELSEN AND ROSENFELD, *ASR*, APRIL 1981)

THOMAS P. WILSON
University of California, Santa Barbara

In a discussion of differential equation models, Nielsen and Rosenfeld (1981:161) argue that "substantive interpretations of dynamic models should focus on the parameters that describe the trajectory, generated by the underlying model of the rate of change, rather than directly on the coefficients of the differential equation." They base this recommendation on what they see as an "intuitive contradiction" (p. 165) between interpretations in the literature for a model of occupational mobility (Sørensen, 1977) and a model of political mobilization (Nielsen, 1980), where the models in question are mathematically equivalent. Nielsen and Rosenfeld then present an "interpretive syntax for differential equation models" and go on to discuss at some length several substantive issues in relation to mobility models and the concept of opportunity. Here I am not concerned with these substantive matters but rather only with the general methodological issues raised by their stricture against interpreting a differential equation directly and their principle of instead interpreting only the coefficients of the solution of the differential equation. My purpose is to suggest that their proposals should be treated with considerable caution.

To begin with, the appearance of contradiction that motivates Nielsen and Rosenfeld's argument results from nothing more profound than ambiguity of terminology. Nielsen and Rosenfeld consider a simplified version of Sørensen's model of individual occupational mobility within a fixed opportunity structure in which the status Y of an individual changes over time in accordance with the differential equation

$$dY/dt = aX + bY,\quad b < 0,\ Y < Y^* \tag{1}$$

where X is a constant over time, though it may vary from individual to individual, and Y^* is the equilibrium value of Y. Nielsen and Rosenfeld (1981:163) summarize Sørensen's interpretation of b as meaning that the smaller $|b|$ is, the greater the opportunity for occupational mobility within the structure. In an-

* Direct all correspondence to Thomas P. Wilson, Department of Sociology, University of California, Santa Barbara, CA 93106.

other study, Nielsen (1978) examined political mobilization and suggested the model

$$dY/dt = -b(Y^* - Y),\quad b < 0,\ Y < Y^*, \tag{2}$$

where Y is amount of electoral support at time t and Y^* is the equilibrium level of electoral support (since $dY/dt = 0$ when $Y = Y^*$). In this case, $|b|$ is interpreted as measuring the speed of mobilization (p. 165). As Nielsen and Rosenfeld observe, the two models are mathematically equivalent: we can find the equilibrium value of Y in Equation 1 by setting $dY/dt = 0$, which gives $Y^* = -(a/b)X$; then, solving for a, we have $a = -Y^*b/X$, and substituting this into Equation 1 gives Equation 2. Moreover, Nielsen and Rosenfeld see contradiction between the two interpretations of b, even though Equations 1 and 2 are mathematically equivalent: "on the one hand, $|b|$ is identified as . . . a measure of the resistance of the system against advancement; on the other as an indicator . . . of lack of resistance of the system against mobilization" (p. 165).

However, this conclusion seems dubious. Without entering into discussion of the merits of either Sørensen's or Nielsen's interpretation, we can note that the substantive contexts of the two interpretations are different, and so it is not obvious that occupational "advancement" or "mobility" is the same or even analogous to political "mobilization," nor is it evident that "resistance" means the same thing in the two contexts. Indeed, the appearance of contradiction evaporates once we realize that Sørensen's interpretation is in terms of how much further occupational advancement is possible for an individual, whereas Nielsen's is in terms of how far a party has come in mobilizing support. Thus, if instead we look at Nielsen's model in terms of how much further support a party can mobilize, or if we view Sørensen's model in terms of how far individuals have come in attaining occupational status, we see that there is no incompatibility. In both cases, the postulated mechanism is one in which Y approaches its equilibrium level Y^* at a rate that declines proportionally with the remaining distance to Y^*, where the constant of proportionality is $-b > 0$. Hence, when b is large in magnitude, the process takes place rapidly, and high political support or occupational status is mobilized or attained in a short time, with little room for improvement in the future; and when b is small in magnitude, the process occurs slowly, and so at any given time there is more room for improvement than when $|b|$ is large. In short, there is no difficulty of the sort imagined by Nielsen and Rosenfeld in interpreting a differential equation substantively.

Thus, the grounds Nielsen and Rosenfeld advance for avoiding direct interpretation of a differential equation are not cogent. Moreover, their proposal to restrict attention to the solution of the differential equation is open to serious objection.

Before turning to the major issue, I mention three relatively minor problems with Nielsen and Rosenfeld's proposal for an "interpretive syntax for differential equation models" (pp. 166–8). First, their discussion applies not to differential equations generally but only to those like Equation 2, the solution of which is

$$Y = e^{bt}Y(0) + (1 - e^{bt})Y^*,\quad b < 0,\ Y < Y^*, \tag{3}$$

where Y(0) is the value of Y when $t = 0$. While equations of this form may be of current interest in the sociological literature, they by no means exhaust the kinds of differential equations that might prove useful in describing social phenomena, nor even do they exhaust the class of linear differential equations with stable equilibria, as indeed Nielsen and Rosenfeld recognize elsewhere in the paper. Consequently, this section should be subtitled more accurately, and appropriate qualifications should appear at several places in the text. Second, the strategy Nielsen and Rosenfeld pursue in this section is of limited generality, for it depends on the differential equation having an explicit solution that can be examined in detail. But it is well known that many differential equations cannot be solved explicitly, and while a numerical solution often can be obtained in a particular application, there is no closed-form expression analogous to Equation 3 that can be examined and interpreted. Here again, a caveat to the reader is in order. Third, Nielsen and Rosenfeld observe that a differential equation can be reparameterized so as, for example, to have the form of Equation 1 rather than Equation 2, and they note (p. 166) that the choice between the two is essentially a substantive one. However, they fail to remark that the solution as well as the differential equation can be reparameterized and instead treat Equation 3 as a fundamental or in some sense most useful form. But, letting $R = Y^* - Y(0)$ be the range of variation of Y in the interval $0 < t < \infty$, we can write the solution of Equation 2 as

$$Y = Y^* - Re^{bt},\quad b < 0,\ Y < Y^*. \tag{4}$$

Clearly, for some substantive purposes, this representation might be far more perspicuous than Equation 3. The importance of these points is that the mathematically unsophisti-

cated reader, to whom this paper is presumably addressed, may well come away from the paper as written with the impression that the discussion in this section is far more general than in fact it is.

However, the major objection to Nielsen and Rosenfeld's proposal to limit substantive interpretations to the solutions of differential equations is that it eliminates what is perhaps the most important application of differential equations in the description of empirical phenomena. Of course, the behavior of a function such as Equation 3 is illuminated by examining the form of the derivative, which in this case is Equation 2; but in this application, the role of the differential equation is completely secondary, and it is only in a trivial and uninteresting sense that one can speak of a differential-equation model. In contrast, a differential equation can be used to represent a causal process that operates, not on the dependent variable directly, but instead on the rate of change of the dependent variable. The problem is that in such a case there is no way of avoiding directly interpreting the coefficients of the differential equation, and that is precisely what we must not do if we follow Nielsen and Rosenfeld. For example, suppose we are studying occupational mobility in a particular industry in which the firms are intensely competitive with one another for personnel. Suppose further that we believe, on the basis of extensive and careful observation and interviews in a number of firms, that the characteristic adaptation of each firm is to design its promotional policies so that individuals are moved up rapidly at first, partly in response to offers from other firms and partly to forestall such offers, but that as individuals acquire some stake in their situations (e.g., relations with coworkers, eligibility for certain fringe benefits, family ties to schools and neighborhoods, etc.), promotions are given less rapidly. To describe this causal process, we might propose Equation 2 as a first approximation. In this case, Y* is the maximum the individual can expect to achieve given his or her line of work, skills, and the like, and b is a measure of competitive pressure the firm is under to promote rapidly during an employee's early years. The crucial point of this story is not whether it is substantively plausible about some particular industry, but rather that it makes plain the fact that we can indeed formulate a causal model in terms of rates of change, which requires a differential or difference equation having a direct substantive interpretation.

The central issue, which Nielsen and Rosenfeld recognize at several points, is that choosing the form of a model that displays the causal structure is fundamentally a substantive matter. But this applies not merely to alternative parameterizations, such as Equations 3 versus 4 or 1 versus 2, but also to whether the causal structure is on the level of the dependent variable itself or on the level of a rate of change.[3] And it is because this point is obscured in their discussion that Nielsen and Rosenfeld's paper is seriously misleading on the general methodological question they address.

REFERENCES

Nielsen, François
1980 "The Flemish movement in Belgium after World War II." American Sociological Review 45:76–94.

Nielsen, François and Rachel A. Rosenfeld
1981 "Substantive interpretations of differential equation models." American Sociological Review 46:159–74.

Sørensen, A.
1977 "The structure of inequality and the process of attainment." American Sociological Review 42:71–91.

INTERPRETING LINEAR DIFFERENTIAL EQUATION MODELS*

(REPLY TO WILSON)

FRANÇOIS NIELSEN
University of Chicago

RACHEL A. ROSENFELD
University of North Carolina–Chapel Hill

The comments made by Wilson on our paper seem to fall into two categories. On the one hand, he reiterates several qualifications and warnings that were already spelled out in the original article, because he thinks these deserve more emphasis. Since, by definition, we generally agree with him on these points, we will discuss them only briefly in this reply. On the other hand, Wilson raises two more important points. First, he suggests that the contradiction we found between interpretations of the coefficients of a linear differential equation in two different substantive contexts results from nothing more profound than ambiguity of terminology. We disagree with Wilson on this point, in large part because we believe ambiguities in terminology to be sometimes more profound than he thinks. We will discuss this comment in detail below. Second, there is Wilson's "major objection" to our paper, which is that a differential equation can be used to rep-

* Direct all correspondence to Rachel A. Rosenfield, Department of Sociology, University of North Carolina, Chapel Hill, North Carolina 27514.

resent causal processes operating not only on the level of some dependent variable, but also directly on the rate of change in this variable. We find this comment puzzling: we wanted to clarify the interpretation of a linear differential equation model precisely because it is a simple model of the rate of change. This particular objection appears to us a result of an incorrect reading of our paper, since the strategy Wilson presents as an alternative is the one we advocate and have used in our own substantive work. Below, we will present evidence for this point as well.

With respect to the first category of comments, Wilson worries about "the mathematically unsophisticated reader," who might be misled into thinking that "the discussion . . . is far more general than in fact it is," and repeats qualifications we had made ourselves. We explicitly restricted our discussion to one of single-equation linear differential equations (of the first order) that have a stable equilibrium (see p. 160 and p. 163, note 5). Wilson is perhaps correct about the need to reiterate this restriction for the careless reader. We certainly agree that the subsection of our paper to which he refers could be retitled from "An interpretive syntax for differential equations models" to "An interpretive syntax for linear differential equation models with stable equilibria." The reader might agree, however, that such rephrasing would have sounded cumbersome. We also knew, before Wilson told us, that many nonlinear differential equations cannot be solved explicitly (p. 160) but are still of interest as mathematical representations of various processes. But linear differential equations always have an explicit solution. Since we had restricted our discussion to linear equations, Wilson's point here is moot. Finally, Wilson notes that the solution (or equation for the trajectory of Y over time) can be reparameterized as a function of the difference between the value of Y at the beginning of the process (Y_0) and at equilibrium (Y^*) (see his Equation 4). He seems to think this refutes our implied (by him) treatment of the particular form of the integral equation as fundamental. We do not see how this point adds to our discussion. Certainly, in our own substantive work we, too, have rearranged the terms of the integral to best fit a particular problem. While Wilson's minor reparameterization is correct (and still allows one to see how the larger $|b|$, the more quickly the "range of variation in Y" is covered), it is not put in the context of a particular substantive problem to demonstrate where it would be more "perspicuous" than our representation.

Most of Wilson's minor comments thus seem to us moot or redundant, at least for readers who, while perhaps "mathematically unsophisticated," are interested enough in our article to read it carefully.

We turn now to the more important comments. In the first of his substantive comments, Wilson dismisses the importance of the contradiction we found in the interpretation of a linear differential equation model in different contexts as mere terminological ambiguity. Strictly speaking, Wilson is correct: our argument ultimately is one of terminology. In contrast with Wilson, though, we do not think this is trivial, for several compelling reasons. First, the ambiguity is serious enough that the same term can mean two opposite things in the same substantive context. The best example is the double meaning of the word "opportunity," the semantic aspects of which we discussed in our paper (p. 171). The coefficient b of the differential equation can be interpreted in terms of the speed of goal attainment on the one hand and in terms of chances for further advancement on the other, a remark Wilson amplifies. The two meanings are contradictory within the dynamic model of achievement. It may be that Coleman's (1968) interpretation of b directly from the differential equation as "negative feedback" is one reason that this ambiguity continued to exist without comment as long as it did. In contrast with Wilson, we obviously believe such a semantic divergence is worth emphasizing, especially since the term "opportunity" is part of the common language, has meanings that have evolved completely independently from differential equations notation, and has emotional and political impact to the point that government agencies have the word in their titles. Second, we have empirical reasons to believe that researchers do transfer their intuitive interpretations of coefficients in particular kinds of models from one substantive context to another. For example, Sørensen and Halliman (1977) transfer interpretation from the context of occupational achievement to that of learning. The problem is that interpretation in one context becomes a paradigm (in the original sense of "template") for interpretation in other contexts. Thus, discussions of terminology are crucial. Finally, we believe that the "substance relative" position Wilson advocates is dangerous for the very logic of scientific inquiry. If interpretations of the coefficients are totally dependent on the substantive context, and two opposite interpretations of the relative magnitude of the same coefficient are possible, then a researcher can always interpret empirical results in such a way as to fit the theory. No hypothesis can be rejected. We thus do not share Wilson's attitude toward this issue.

The second substantive point made by Wil-

son, which he considers to be the major one of his comment, is that by focusing on the integral equation, we ignore the fact that ". . . a differential equation can be used to represent a causal process that operates, not on the dependent variable directly, but instead on the rate of change of the dependent variable." We find it difficult to reply to such a remark because the analysis of the rate of change of a dependent variable, as opposed to the level of a variable at some time, is exactly what we advocate, not only in principle but also in practice. Rosenfeld (1980), for example, explains variations in rates of income and status mobility among groups by differences in the structure of labor markets faced by these groups, while Nielsen (1980) speculates about the effects of parties' organizational structures on the speed of political mobilization or demobilization. Our interest is in change over time, not levels of outcomes, except insofar as we are also interested in Y*, the equilibrium level. The only way we can explain Wilson's comment is that he was impressed by our emphasis on the equation of the trajectory, which we used as an expository device to show concrete implications of a dynamic model—i.e., how levels of Y change over time at different speeds—to the point of missing our reason for writing the article in the first place—i.e., our concern that dynamic models be correctly used to model processes, or rates of change.

To summarize, with one major exception, we find it difficult to disagree with Wilson's comments, since he is simply reiterating what we have already said. We appreciate him having taken the time to reemphasize these points for *ASR* readers, although we are troubled by the interpretations he imputes to us that are opposite the ones we hold, especially that with respect to the use of differential equation models to explain rates of change. We do disagree with his contention that interpretation of coefficients depends entirely on the substantive context. To conclude, Wilson has given us no reason not to stand by our story.

REFERENCES

Coleman, James S.
1968 "The mathematical study of change." Pp. 428–78 in H. M. Blalock and A. B. Blalock (eds.), Methodology in Social Research. N.Y.: McGraw-Hill.

Nielsen, François
1980 "The Flemish movement in Belgium after World War II." American Sociological Review 45:76–94.

Nielsen, François and Rachel A. Rosenfeld
1981 "Substantive interpretations of differential equation models." American Sociological Review 46:159–74.

Rosenfeld, Rachel A.
1980 "Race and sex differences in career dynamics." American Sociological Review 45:583–609.

Sørensen, Aage B. and Maureen Halliman
1977 "A reconceptualization of school effects." Sociology of Education 50:273–89.

ISSUES IN INSTRUMENTAL VARIABLES ANALYSIS*

(COMMENT ON CRAMER, *ASR*, APRIL 1980)

KARL L. ALEXANDER
THOMAS W. REILLY
JAMES FENNESSEY
The Johns Hopkins University

In a recent article in this journal, Cramer (1980) undertook to reconcile discrepant conclusions from two studies of the relationship between fertility and female employment. Both studies he cites (Smith-Lovin and Tickamyer, 1978; Waite and Stolzenberg, 1976) used nonrecursive structural equation models. Cramer criticizes this analysis technique by pointing out that the choice of instrumental variables in nonrecursive models often is problematic (see also Land and Felson, 1978) and that discrepancies across research studies using such techniques may result from the properties of the instrumental variables used. Cramer presents the results of several consistency analyses examining the use of alternate instrumental variables in 2SLS estimation, and shows little consistency in the estimates produced. He also introduces the M^2 statistic as a diagnostic tool for distinguishing weak from strong instruments in 2SLS analysis. This statistic quantifies the redundancy of the constructed endogenous predictor in the second-stage matrix with the other predetermined variables from the equation in which it appears. He focuses on this "collinearity" in the second-stage matrix of 2SLS estimation as the basis for diagnosing weak instruments.

We believe Cramer's preoccupation with multicollinearity in evaluating instrumental variables is unduly narrow and that it deflects attention from the broader issues of model specification and theoretical adequacy that are more fundamental to the suitability of a given instrumental variable. The purpose of this

* Direct all correspondence to Karl L. Alexander, Department of Social Relations, The Johns Hopkins University, Baltimore, Maryland 21218.

This research was supported in part through two NIMH research training programs in the Department of Social Relations: 1 T32 MH15735 and 2 T32 MH14587. The opinions expressed herein are those of the authors and not necessarily those of NIMH.

comment is to demonstrate the limited usefulness of M^2 as a diagnostic in 2SLS analysis. First, we describe some of the technical properties of Cramer's M^2 statistic that might not be apparent from his original exposition. Second, we provide a concrete example of discrepant instrumental variable estimations for which Cramer's collinearity criterion offers little guidance. And, third, we review some of the broader considerations in the evaluation of any instrumental variable analysis.

TOOLS FOR ASSESSING ESTIMATION PRECISION

The technique of consistency checks—using several alternative instruments separately and evaluating the similarity in results—is valuable, and Cramer has employed it constructively. When inconsistencies do appear, any of several problems may be at issue: substantive specification errors, errors in the assumptions made to achieve identification, or statistically imprecise instruments.

Only after the issues of errors in specification and identification assumptions have been resolved is it appropriate to focus on the matter of precision of estimates, as Cramer does in his consistency analysis. Cramer's presentation tends to equate precision with the "collinearity" measured by his M^2 statistic. But the precision of an estimate (the reciprocal of the variance) depends as well on the covariance between the included endogenous predictor (the "instrumand," denoted here as Y_2), and the variable used as an instrument for it (denoted as $\hat{Y}_2$), net of the included predetermined variables. Since the variance of $\hat{Y}_2$ is equal to its covariance with Y_2, its precision, in fact, depends on the product

$$V[\hat{Y}_2](1-M^2) = \mathrm{Cov}[Y_2\ \hat{Y}_2](1-M^2).$$

Because Cramer's presentation neglects the first term in these products, it is incomplete as a measure of the precision of an estimate from an instrument. This incompleteness also belies Cramer's statement (p. 170) that the M^2 statistic is inversely proportional to the standard error of the coefficient. On the other hand, if this statement were true, then the M^2 statistic would simply duplicate an existing measure of precision of estimate, the standard error. Moreover, as a measure of collinearity, the M^2 statistic is equivalent to $(1 - \text{tol})$, where "tol" is the tolerance value reported in the SPSS regression package and in the BMD regression program. In summary, Cramer's M^2 statistic is identical to an existing measure of collinearity and lacks the exactness of the standard error as an index of the statistical quality of an estimator. On these technical grounds, then, its introduction would seem to add little to extant criteria for evaluating the precision of an estimator and, therefore, to our ability to diagnose inadequate instruments.

ANOTHER EXAMPLE OF ALTERNATIVE INSTRUMENTS

From our own experience in modeling the relationship between age at marriage and educational attainment (Alexander, Reilly, and Eckland, 1979; Alexander, Reilly, and Fennessey, 1980), which parallels closely Cramer's exposition, we too were struck by the differences between studies: our own nonrecursive results differed considerably from those obtained by others using conceptually similar simultaneous-equations approaches (Marini, 1978; Voss, 1977). We suspected, as did Cramer, that these discrepancies might be due to differences in the properties of the instrumental variables used in these studies. To examine this possibility, we too performed a consistency analysis of alternative instrumental variables within our own data set (see Alexander, Reilly, and Fennessey, 1980). Our choice of candidate instruments was informed by the specifications used by Marini and by Voss, although we were unable to reproduce some of their measures as closely as we would have liked.

Our instrumental variables analysis most pertinent to Cramer's discussion of collinearity in 2SLS estimation is presented in Table 1. Five alternative instrumental variables are used to identify the effect of age at marriage on educational attainments. These are:

- SPOUSE: a dichotomous measure which indexes whether the respondent met his/her spouse while in high school.
- RETARD: measures the respondent's age at high-school graduation.
- AGE: measures the respondent's age during his/her sophomore year in high school.
- DATE: sums the scores on two survey questions referring to the respondent's high-school peer groups: how important dating was for being popular among one's closest friends, and whether the respondent belongs to the "party-going" group.
- SEX: indexes whether the respondent had "sex relations" while in high school with anyone other than his/her eventual spouse.

Two alternative instruments are used to identify the effect of educational attainment on age at marriage:

- EDEX: indexes the respondent's educational expectations as expressed during high school.

Table 1. Estimates of the Reciprocal Relations between Educational Attainment (EDUC) and Age at First Marriage (AGEMAR), Based on Use of Alternative Instrumental Variables

Dependent Variable	Reciprocal Regressor[a]	Instrument	Effect of Reciprocal Regressor on Dependent Variable: Metric	Standardized	Standard Error	M^2
		Males (N = 553)				
EDUC	AGEMAR	SPOUSE	.413	.339*	.095	.293
EDUC	AGEMAR	RETARD	−4.760	−3.903	7.860	.992
EDUC	AGEMAR	AGE	−1.415	−1.161	.748	.887
EDUC	AGEMAR	DATE	−.480	−.394	1.016	.973
EDUC	AGEMAR	SEX	−.237	−.195	.281	.767
AGEMAR	EDUC	EDEX	.176	.215	.140	.788
AGEMAR	EDUC	CURRIC	.287	.350*	.129	.753
		Females (N = 615)				
EDUC	AGEMAR	SPOUSE	.388	.409*	.137	.634
EDUC	AGEMAR	RETARD	7.511	7.911	33.892	.999
EDUC	AGEMAR	AGE	−.457	−.481	.425	.886
EDUC	AGEMAR	DATE	56.878	59.910	1935.523	1.000
EDUC	AGEMAR	SEX	3.008	3.169	8.446	.998
AGEMAR	EDUC	EDEX	.239	.227	.164	.864
AGEMAR	EDUC	CURRIC	.586	.556*	.116	.767

[a] Predetermined predictors in the EDUC equation: EDEX, mother's education, father's education, an index of household possessions, father's occupation, academic aptitude. Predetermined predictors in the AGEMAR equation: SPOUSE, mother's education, father's education, and index of household possessions, father's occupation and academic aptitude. See Alexander, Reilly, and Eckland, 1979, for details on measurement.

* Coefficient more than twice its standard error.

CURRIC: a dichotomous measure which identifies the respondent's high-school curriculum membership, contrasting college-preparatory enrollment with all others.

The other predictor variables in each equation are several measures of socioeconomic background and scores from a standardized test of academic ability.

Our data are from the EEO (Explorations in Equality of Opportunity) survey of 1955 high-school sophomores who were resurveyed as young adults (about age 30) in 1970. The sample was drawn from 42 high schools scattered throughout the United States. Usable follow-up data were obtained from 1,130 women and 947 men. Our analyses are based upon the ever-married subsamples of these youth for whom data were available on all of the measures used in the regressions. The sample sizes for the analyses are 615 for women and 553 for men. See Alexander, Reilly, and Eckland, 1979; Alexander and Reilly, 1981; and Alexander, Reilly, and Fennessey, 1980.

Many of the instrument candidates produce transparently implausible results in the coefficients of reciprocal influence: standardized coefficients are far above unity, and standard errors of estimation are very large (see Table 1). In these instances, then, the results themselves indicate that something probably is amiss. For the effect of age at marriage on educational attainment, SPOUSE is the only instrument that produces results at all reasonable. When this instrument is used, a significant, sizable positive effect is obtained for both men and women. The estimations obtained from RETARD, AGE, DATE and SEX all are either opposite in sign from what would be expected or excessively large. None, moreover, is statistically significant at conventional levels. In these instances the patterning of the M^2 coefficients is reasonably consistent with the clues provided by the results. With one exception, the M^2 coefficients associated with the structural estimations deemed either out of range or implausible are all above .85, and most are well above .90. Although there is no absolute standard to guide such judgments, collinearity in this range generally would be considered cause for caution. On the other hand, the M^2 statistics obtained when SPOUSE is the instrument are both noticably lower: .293 for men and .634 for women.

At least in these instances Cramer's diagnostic points in the same direction as other indications of potential trouble. For these comparisons, then, M^2 might serve as a useful heuristic. Such is not as clearly the case, however, with the analysis performed on the alternative instruments for educational

attainment—educational expectations and curriculum placement (see Table 1).

When curriculum and educational plans are used alternatively to identify the coefficient expressing the effect of educational attainment on age at marriage, markedly different estimations are obtained, though in this instance none stands out immediately as unreasonable. For both men and women, the estimates derived when curriculum enrollment is the instrument are large and statistically significant. Those obtained when educational plans is the instrument are smaller and nonsignificant. Hence, when curriculum is used as an instrument we would conclude that educational attainment is an important factor in marriage timing; when educational plans are used, however, we would conclude that the effect of educational level on marriage timing is insignificant.

Even if significance levels were ignored, the absolute differences in coefficients are still substantial. Among women, the effect obtained by using curriculum is more than twice that obtained by using educational plans (b's of .556 and .227). For men, the values are .350 and .215, a smaller but still notable difference (although the gap here is less than twice either's standard error).

What makes these differences relevant in the present context is that none of the diagnostics discussed thus far provides a basis for making sense of them. The results are all plausible, though quite different in implication. Moreover, though Cramer's M^2 statistic might suggest discounting the results for educational plans as an instrument among women (.864), the M^2 values are almost identical among men.

DISCUSSION

In evaluating a structural equation analysis, we think it useful to distinguish four logically ordered levels of concern: first, the broad causal imagery most appropriate to the issues; second, the conceptual specification of the research model—enumerating the relevant variables and the influence flows among them; third, identification—selecting and justifying suitable instruments; fourth, checking the statistical quality of the estimates. These levels are, in some sense, nested, but Cramer's discussion does not reflect a sensitivity to this implicit hierarchy. Cramer's consistency analyses reveal little consistency across instruments, suggesting to him some sort of specification problem. He then shifts to the level of broad causal imagery, arguing that questions regarding patterns of life-course development are more properly studied in a dynamic longitudinal framework than via the static, nonrecursive approach adopted by both Smith-Lovin-Tickamyer and Waite-Stolzenberg.

Although the failures of Cramer's consistency checks do indicate that something is wrong, they do not imply that the difficulties are at the level of broad causal imagery. His discussion leap-frogs from considerations at the fourth level of the hierarchy to considerations at the first level, ignoring almost entirely issues of model specification generally and assumptions regarding instrumental variables particularly. Cramer's preference for dynamic models is not ill-founded, but the switch to a dynamic approach neglects issues that arise at the second and third levels of concern. By focusing his attention on the empirical quality of the instruments (i.e., multicollinearity) and then shifting to concerns of broad causal imagery, he fails to consider questions of model specification and theoretical adequacy within an instrumental variables framework.

Cramer summarizes the implications of his critique of instrumental variables analysis as follows:

> Nonrecursive models are not cheap devices for easily determining causal direction. The main problem, aside from specifying all relevant variables (a problem in all models), is weak instrumental variables, or multicollinearity. . . . The theoretical challenge of identifying instrumental variables is matched, especially with microlevel data, by the empirical problem of finding good instruments. If instruments are weak, the "medicine" of 2SLS may be worse that the "disease" of unclear causal direction. [Cramer, 1980:187–8]

This statement seems to imply that nonrecursive models are suitable devices for determining causal direction, cheaply or otherwise. However, such analyses quantify the implications of a specific set of assumptions regarding causal processes: a subtle yet important difference from determining causal direction. In addition the statement limits the theoretical burden of model construction to the task of identifying all relevant variables. This, though, is only part of the burden, the other part being to specify the relationships among these variables. The statement also equates the empirical problem of finding good instruments with the collinearity property indexed by Cramer's M^2 statistic, i.e., that the instruments not be "weak" by his standard. As noted above, the M^2 statistic is incomplete as an empirical diagnostic. More generally, preoccupation with the strength of instrument candidates neglects many other considerations, both empirical and theoretical, that determine the usefulness of a particular instrument for a particular purpose. Low collinearity may well be a necessary quality of a good instrument, but it alone is hardly sufficient.

Gillespie and Fox (1980) have suggested a different diagnostic for potential problems in simultaneous equations models—the appearance of negatively correlated structural disturbances where positive covariance would be anticipated. They demonstrate, both analytically and with simulated data, some of the general conditions that might give rise to such anomalies: positive covariance between instrumental variables and equation disturbances, inappropriate a priori structural constraints (i.e., assumed zero effects), neglected sources of spuriousness in the instrument–endogenous variable relationship, and random and nonrandom errors in the measurement of exogenous variables.

Although Gillespie and Fox's discussion is not exclusively in terms of instrumental variables, all of their examples locate these problems either in the properties of instrumental variables or in the relationships between instrumental and endogenous variables. Their discussion more realistically conveys the demanding theoretical conditions that must be satisfied if instrumental variables estimations are to be believable than does Cramer's collinearity criterion.

Sometimes empirical diagnostics such as those recommended by Cramer and by Gillespie and Fox encourage researchers to further reflection and troubleshooting. Sometimes, too, recourse to data-based correctives such as consistency checks and sensitivity analyses suggests which measures and/or assumptions are serviceable. Most often, though, these sorts of empirical checks are only suggestive, and for many concerns data-based guidance is lacking entirely. As Namboodiri, Carter, and Blalock note (1975:552–5), many such issues must rest upon the researcher's judgment.

We are in wholehearted agreement with the sentiment of Cramer's concluding arguments. Convincing nonrecursive estimations do not come cheaply. Similarly, simply seeking out "strong" instruments is not a cheap empirical fix. At their root, such models are driven by a demanding set of untestable assumptions regarding influence flows and, often, equally severe assumptions regarding measurement properties. Diagnostics such as multicollinearity and negatively correlated disturbances sometimes suggest deficiencies of data or theory, and sometimes there may be sufficient information available to check out the plausibility or implausibility of the more tenuous assumptions, but these signals from the data are rarely definitive in themselves. Convincing causal models are not data-driven. Rather, what makes them convincing is the cogency of the theory and thinking reflected in their specification. Results may suggest when a model is not credible, but they alone cannot tell us when a model is credible. Contrary to Cramer, the main problem with instrumental variables analysis is more often weak theory than weak instruments. In the absence of a convincing conceptualization, concerns regarding the empirical properties of candidate instruments are moot.

REFERENCES

Alexander, Karl L. and Thomas W. Reilly
1981 "Estimating the effects of marriage timing on educational attainment: some procedural issues and substantive clarifications." American Journal of Sociology 87:143–56.

Alexander, Karl L., Thomas W. Reilly, and Bruce K. Eckland
1979 "Family formation and educational attainment." Paper presented at the Research Conference on Educational Environments and the Undergraduate Woman, Wellesley, Mass.

Alexander, Karl L., Thomas W. Reilly, and James Fennessey
1980 "Tuning the instruments: modeling age at marriage and educational attainment." Technical report, Department of Social Relations, Johns Hopkins University: Baltimore, MD.

Cramer, James C.
1980 "Fertility and female employment: problems of causal direction." American Sociological Review 45:167–90.

Gillespie, Michael W. and John Fox
1980 "Specification error and negatively correlated disturbances in 'parallel' simultaneous equation models." Sociological Methods and Research 8:273–308.

Land, Kenneth C. and Marcus Felson
1978 "Sensitivity analysis of arbitrarily identified simultaneous equation models." Sociological Methods and Research 6:283–307.

Marini, Margaret M.
1978 "The transition to adulthood: sex differences in educational attainment and age at marriage." American Sociological Review 43:483–507.

Namboodiri, N. Krishnan, Lewis F. Carter, and Hubert M. Blalock
1975 Applied Multivariate Analysis and Experimental Designs. New York: McGraw–Hill.

Smith-Lovin, Lynn and Ann R. Tickamyer
1978 "Nonrecursive models of labor force participation, fertility behavior, and sex role attitudes." American Sociological Review 43:541–57.

Voss, Paul R.
1977 "Social determinants of age at first marriage in the United States." Paper presented at the Annual Meetings of the Population Association of America, St. Louis, Missouri.

Waite, Linda and Ross Stolzenberg
1976 "Intended childbearing and labor force participation of young women: insights from nonrecursive models." American Sociological Review 41:235–52.

INSTRUMENTAL VARIABLES ANALYSIS

(REPLY TO ALEXANDER ET AL.)

James C. Cramer
University of California at Davis

The comment by Alexander et al. more clearly resembles an affirmation and extension of my paper than a critique. The main points of their criticism are that M^2 is not a sufficient statistical diagnostic and that statistical problems in nonrecursive models are emphasized to the exclusion of theoretical problems. Both points are based on distortions of my article; in fact, we essentially agree on both points.

It is clearly stated in my paper that M^2 is a new notation but hardly a new statistical concept. Equivalent statistics are discussed in a number of texts and software manuals, as noted by Alexander et al. and in my paper. M^2 is algebraically related to the standard error of the regression coefficient for the endogenous predictor, but the two statistics serve different purposes. The standard error is useful for statistical inference, but it has no upper bound and is not useful as a descriptive statistic. M^2 is standardized to the range (0,1), so it is a useful description of the degree of multicollinearity. That is all M^2 is. Alexander et al. seem to desire more: a numerical indicator of whether an estimated coefficient is "correct" or not. High multicollinearity means that there is a good chance that an estimated coefficient is "incorrect;" yet it is quite possible to obtain a good estimate despite a high value of M^2 and to obtain a bad estimate despite a low value of M^2. Implausible estimates actually are better predicted by M^2 in their sensitivity tests than in mine, but Alexander et al., because of unreasonable expectations, are more dissatisfied with M^2. The lesson is to use M^2 as evidence (but not proof) of the quality of a model, along with any other evidence available, and always to exercise judgment.

At several places in my article, it is clearly stated that strong theory and proper model specification are essential in quantitative research. The necessity of good theory does not mean that all other considerations are trivial or irrelevant, and it is not improper to write about other considerations. Theoretical problems of specification and identification in nonrecursive models are widely discussed in texts and research reports and are well-known to practitioners. Except for the common, bland warning that two-state least squares estimates are "consistent but inefficient," statistical problems with nonrecursive models are seldom discussed. In my field of research, theoretical problems are reasonably adequately discussed but statistical problems are not, and my paper was addressed partly to this imbalance. Theoretical problems, however, certainly were not neglected. Contrary to Alexander et al.'s assertion, definitions and measures of variables and relationships between variables are discussed, as well as the "causal imagery" of dynamic and static formulations. Indeed, the main concern of the paper was theory, and Alexander et al. aim their criticisms at a secondary part of the argument.

Although Alexander et al. strongly emphasize the theoretical aspects of nonrecursive models, it is clear that they do not entirely disregard the statistical aspects. If they believed entirely in theory, they would not perform sensitivity analyses; for if theory is strong, then each model in a sensitivity analysis must be misspecified and the results misleading (due either to omitted variables or low reliability). In my paper, sensitivity tests were justified (implicitly) on the grounds that the underlying theory, although adequately discussed, is not terribly strong. That is, compelling reasons do not exist for preferring one or another of the several extant models. Evidently Alexander et al. find themselves in a similar situation. I believe that this situation (i.e., weak theory) is, in fact, quite common.

Furthermore, in their research, the instrument for age at marriage that is preferred by Alexander et al., on statistical grounds, seems to be one of the least defensible theoretically. Many people date in high school. Whether a high-school friend becomes a spouse is determined at marriage, not before (regardless of intentions or formal vows, which may change), and probably depends partly on the interval from high school to marriage—i.e., essentially, on age at marriage. In addition, this specification seems to suggest that age at marriage depends on meeting the right person, so it is uninteresting as well as circular. Clearly Alexander et al. are not guided solely by theory.

My models too are somewhat circular and uninteresting; my point is not to criticize Alexander et al. Theory is essential in reducing the number of plausible models. In many areas of survey research, however, theories are not yet strong enough to synthesize past research and speculation into a single model. Statistical considerations, such as multicollinearity, often may have a useful role in model building. As we see in Alexander et al.'s research, theoretically plausible variables may be deficient statistically, and statistically satisfactory variables may be deficient theoretically. Both criteria must be considered. M^2 is a useful but neglected guide in this stage of research, but by itself is not sufficient. Thus I believe that Alexander et al. and I are in basic agreement.

SOME USEFUL APPLICATIONS OF LOGISTIC MODELS: REPLY TO AN ODD CRITIQUE

(COMMENT ON SWAFFORD, *ASR*, AUGUST 1980)

STEVEN L. GORTMAKER
Harvard School of Public Health

A recent ASR article (Swafford, 1980) presents an overview and critique of parametric techniques for the analysis of contingency tables. An incorrect assertion appears in the Swafford article, and a very important property of logistic models for the analysis of discrete multivariate data is ignored.

In a section discussing "eight common errors" made in the application of these techniques, Swafford begins by asserting that

> . . . errors are by no means confined to early sociological applications. Gortmaker (1979:286) claims that the criterion variable of his logit model is the log of the death *rate* instead of the log of the *odds* of dying rather than surviving; he misinterprets his coefficients accordingly. [Emphasis from the original.]

This assertion is wrong. I note quite clearly in my article that the sampling plan consists of "probability samples of one out of every 1,000 births and one out of every 110 infant deaths" (Gortmaker, 1979:286). The dependent variables in the logit models estimated thus consist of deaths during a particular time period divided by live births during the same time period. "The conventional infant mortality rate is defined as the number of infant deaths per year per 1,000 live births during the year," and neonatal and postneonatal mortality rates are similarly defined with reference to the deaths in a given period per 1,000 live births during the year (Shryock and Siegel, 1973:410–1). Thus, the models estimated in my article are predicting a dependent variable which consists of infant deaths divided by live births, and the dependent variable is quite rightly referred to as an expected death rate. Swafford is wrong in referring to this dependent variable as the odds of surviving rather than dying.

Swafford apparently assumed that the data being analyzed consisted of a sample of births which were then cross-classified by whether or not a child survived during the ensuing year. In a situation such as this, the dependent variable would be the odds of dying rather than surviving. By not taking into account the sampling plan of my analysis, Swafford thus ironically is ignoring one of his own stated criteria for "good" data analysis (1980:685).

The implications of this incorrect comment, however, extend beyond this single analysis. Swafford has also ignored a number of extremely important applications of logistic and log-linear models which can be usefully applied to just such differentially sampled data. In a variety of situations these models could have much to recommend them over other analyses which might be done employing a linear models approach.

Before discussing some of these applications, however, it is useful to note some important properties of a parameter which is often used in characterizing a cross-classification: this is the odds ratio (see e.g., Mosteller, 1968; Bishop et al., 1975). First, it has often been demonstrated that the odds ratio is invariant to row and column multiplications by a constant (Mosteller, 1968; Bishop et al., 1975). Thus, the different categories of either of two cross-classified variables can be differentially sampled (i.e., a different sampling fraction can be used), but the estimated odds ratio will not be biased. It has also long been recognized that as odds become small and approximate rates, odds ratios approximate estimates of relative risk (Cornfield, 1951). Finally, a variety of authors have noted how similar estimates of odd ratios can be obtained from both retrospective and prospective designs (e.g., Fienberg, 1977:106).

These properties are applied extensively in the epidemiological literature, and indeed are responsible for making possible much of the multivariate analysis of "rare" events via retrospective and case-control studies (e.g. Breslow et al., 1979). In particular, analyses of the etiology of rare causes of death would be severely impaired if these properties were not exploited—thousands of individuals would have to be sampled and followed-up if such rare events could only be studied in the manner implied by a prospective design which analyzed the "odds of dying rather than surviving."

Logistic models are extremely useful in the analysis of such data because the coefficient estimates can be interpreted in terms of relative risks. One of the interesting points of my paper, I thought, was an illustration of how this interpretation could be used with such differentially sampled data. There are certainly many other applications in sociology where these properties can also be exploited. For example, studies of the attainment of rare events, and of the relative risks connecting other variables to the attainment of these events, could be designed around these techniques. Elite membership, the accumulation of great wealth, or the attainment of other statuses could all be studied retrospectively collecting both random samples of these individuals as well as random samples of all other individuals. It would obviously be much more expensive to study the attainment of elite

status, for example, by sampling the total United States population and then looking at the "odds of elite attainment versus not." The sample size required for such a study could be ridiculously large.

REFERENCES

Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul Holland
1975 Discrete Multivariate Analysis: Theory and Practice. Cambridge, Massachusetts: MIT Press.

Breslow, N. E., N. E. Day, K. T. Halvorsen, R. L. Prentice, and C. Sabai
1978 "Estimation of multiple relative risk functions in matched case-control studies." American Journal of Epidemiology 108:299–307.

Cornfield, J.
1951 "A method of estimating comparative rates from clinical data, applications to cancer of the lung, breast, and cervix." Journal of the National Cancer Institute 11:1269–75.

Fienberg, Stephen E.
1977 The Analysis of Cross-Classified Categorical Data. Cambridge, Massachusetts: MIT Press.

Gortmaker, Steven L.
1979 "Poverty and infant mortality in the United States." American Sociological Review 44:280–97.

Mosteller, F.
1968 "Association and estimation in contingency tables." Journal of the American Statistical Association 63:1–28.

Shryock, Henry S. and Jacob S. Siegel
1973 The Methods and Materials of Demography, Volume 2. Washington: U. S. Department of Commerce, Bureau of the Census.

Swafford, Michael
1980 "Three parametric techniques for contingency table analysis: a nontechnical commentary." American Sociological Review 45:664–90.

MORTALITY RATES IN A TWO-SAMPLE STUDY*

(REPLY TO GORTMAKER)

MICHAEL SWAFFORD
Vanderbilt University

I welcome this opportunity to elaborate on Gortmaker's article. Obviously, Gortmaker is still convinced that his criterion variable was a mortality rate.[1] However, the ratio which constituted his criterion variable was not infant deaths divided by live births, but one out of 110 infant deaths divided by one out of 1,000 live births. Since births and deaths were sampled in differing proportions, the ratio does not qualify as a rate. For example, Table 1 in Gortmaker's article (1979:284) reports that his case base included 188 neonatal deaths and 907 live births in circumstances of extreme poverty. The ratio of these two frequencies would provide an outrageous estimate of the neonatal death rate even under circumstances of extreme poverty: 207 per 1,000 live births. The correct figure, reported in the same table, is 22.8 per 1,000. In this particular case, it would be simple to weight the observed frequencies so that the ratio would yield a rate. However, Gortmaker gives no indication that he did so, and in any case, tampering with the observed frequencies would ruin tests of significance unless special adjustments were made.

Table 1. Relative Risk and Odds Ratios in a Two-Sample Study

	Poverty	Other
	Population	
(1) Deaths	60	80
(2) Survivors	1,940	3,920
(3) Births	2,000	4,000
	Two Samples[a]	
(4) Deaths	30	40
(5) Survivors	194	392
(6) Births	200	400

[a] The frequencies in the lower half of the table were obtained by sampling half of the deaths in the population (line 1) and one-tenth of survivors and of births in the population (lines 2 and 3). (For the sake of clarity, no sampling error has been introduced.) Deaths and survivors do not sum to births in the lower half because deaths were sampled in different proportions than survivors and births.

In light of this error, and in the absence of a clear description of Gortmaker's statistical methodology in his original article, readers in search of further detail might well have resorted to the methodological works cited in his article. These, however, give rise to more questions about Gortmaker's approach than they answer. The articles by Goodman do not concern two-sample studies, and they contain no equations in which the criterion variable is a rate. As for Bishop and Mosteller's contribution to *The National Halothane Study* (1969:238–72), although their analysis did yield mortality rates under various kinds of anesthesia, it did so by splitting the table into two parts (pertaining to those undergoing surgery and those dying after surgery), modeling frequencies (not rates) in each of the two parts,

* Address all communications to Michael Swafford, Department of Sociology and Anthropology, Vanderbilt University, Nashville, Tennessee 37235.

[1] My reasons for preferring "criterion variable" to "dependent variable" in this context is given in fn. 3 of my article (1980:666).

and substituting the "smoothed" frequencies into an equation for rates. Bishop, in fact, has denied that she was modeling logs of rates with log-linear equations.[2] In a related article, although she repeatedly refers to tables of rates, she speaks of fitting a linear model to the logit of the rates, not the log of the rates (1969:383, 391). Nor does Gortmaker's reference to Bishop et al. (1975) clarify matters. Although this text contains hundreds of references to at least 80 examples, the only directly relevant example of the two-sample approach seems to be the text's treatment of the halothane study just disqualified as an example of his approach. While reading through the text, however, one is struck by at least two potential problems with Gortmaker's approach. First, the authors admonish readers not to confuse odds ratios with relative risks (p. 15), yet Gortmaker uses a program for the former while couching the discussion in terms of the latter without explaining why. Also, although the authors discuss several sampling designs which produce tables amenable to log-linear analysis, Gortmaker's design is conspicuously absent. The three common sampling distributions which yield tables amenable to log-linear analysis are the multinomial, product-multinomial and independent Poisson distributions (Bishop et al., 1975:62–4). Gortmaker's table does not exhibit any of these distributions because the categories "born" and "died" are not mutually exclusive (Feller, 1950:116, 118, 124). Admittedly, mathematical statisticians are hard at work adapting log-linear methodology to other designs. However, his original article does not cite their work, and his reference to Breslow et al. in his comment actually refers to matched retrospective designs.

In short, Gortmaker erred in describing his criterion variable and miscued readers by referring to methodological works which do not provide obvious support for his tack in conducting the analysis.

Fortunately, it turns out that Gortmaker's erroneous claim about his criterion variable is not crucial to his argument, as is illustrated by the fictitious figures in Table 1. In the population (lines 1–3), 60 of 2,000 poverty births ended in infant deaths (30 per 1,000), while 80 of 4,000 other births ended in this manner (20 per 1,000). Obviously, infants born into poverty were 1.5 times as likely as others to die. This figure can be calculated by applying the formula for the cross-product ratio: $1.5 = (60 \times 4{,}000)/(80 \times 2{,}000)$. Now suppose that half of all deaths and one-tenth of all births were sampled. Disregarding sampling error, the figures in lines 4 and 6 would be obtained. As is the case with Gortmaker's data, the ratio of deaths to births in this section of the table does not yield a rate. Nevertheless, the cross-product ratio again yields the correct relative risk. That is, $(30 \times 400)/(40 \times 200) = 1.5$. Furthermore, even if we did not know what proportion of births and deaths were sampled and could not, therefore, reconstitute the absolute population mortality rates, the cross-product ratio would still correctly reflect the fact that children in poverty were 1.5 times as likely as others to die.

Since cross-product ratios figure heavily in log-linear models, it would seem convenient to employ log-linear models as Gortmaker has. Unfortunately, since births and deaths are not mutually exclusive categories, the assumptions underlying the maximum-likelihood estimates are probably violated. However, when the mortality rate is quite small, the violations are no doubt inconsequential, and the difference between odd ratios and relative risks are miniscule. For example, if we were to employ data for the mutually exclusive categories of deaths and survivors in Table 1, the proper odds ratio would be $(30 \times 392)/(194 \times 40) = 1.52$—a figure quite close to the relative risk calculated in the previous paragraph.

Even though Gortmaker seems confused about his criterion variable, it is clear from his comment that this is his justification for the statistical approach he employed. This straightforward justification is absent from his original article and, if I am not mistaken, from the methodological studies he cites in it. It is unfortunate that the virtues of his tack were obscured by his erroneous claim about the criterion variable and by other miscues in his exposition. I am pleased, however, that the confusion wrought by these facts has been brought to light, for I agree with him that two-sample methodology has much to offer sociologists.

I should like to clarify a statement in my article in which I take issue with Davis for claiming that log-linear modeling assumes a simple random sample and criticize him for failing to constrain his models to reflect NORC's sampling design (Swafford, 1980:685). I was correct in stating that simple random sampling is not assumed in log-linear modeling. However, as Davis has pointed out to me in a gracious letter, NORC's sampling design was not merely stratified, but clustered. Thus, while he could have constrained his models to reflect stratification, the tests of significance produced by ECTA (like those produced by almost every statistical package) would still fail to take into account clustering.

[2] Personal communication, August 12, 1981.

REFERENCES

Bishop, Yvonne M. M.
1969 "Full contingency tables, logits, and split contingency tables." Biometrics 25:119–28.

Bishop, Yvonne M. M. and Frederick Mosteller
1969 "Smoothed contingency-table analysis." Pp. 238–72 in John P. Bunker, William H. Forrest, Jr., Frederick Mosteller, and Leroy D. Vandam (eds.), The National Halothane Study. Washington, D.C.: U.S. Government Printing Office.

Bishop, Yvonne M. M., Stephen E. Fienberg and Paul W. Holland
1975 Discrete Multivariate Analysis: Theory and Practice. Cambridge, Mass.: MIT Press.

Feller, William
1950 An Introduction to Probability Theory and Its Applications. Vol. 1. New York: John Wiley.

Gortmaker, Steven L.
1979 "Poverty and infant mortality in the United States." American Sociological Review 44:280–97.

Swafford, Michael
1980 "Three parametric techniques for contingency table analysis: a nontechnical commentary." American Sociological Review 45:664–90.

LOGISTIC REGRESSION MODELS WITH CONTINUOUS INDEPENDENT VARIABLES*

(COMMENT ON SWAFFORD, *ASR*, AUGUST 1980)

PAUL P. L. CHEUNG
HERBERT L. SMITH
University of Michigan

In an otherwise informative article on the analysis of discrete dependent variables, Swafford (1980) makes some misleading statements about the presumed inutility of logistic regression (logit) models in the presence of continuous independent variables. In particular (p. 686):

> First, when one has a proper categoric dependent variable but continuous independent variables, one ought to consider discriminant analysis (readily available in SPSS). If one's independent variables include both continuous and categoric ones, there may be no fully satisfactory technique. If there happen to be enough cases at each level of the continuous variable, one may utilize the GSK procedure and assign numeric values to each level of the variable. Otherwise, one may be compelled to include categoric variables in a discriminant analysis or to categorize continuous variables for the sake of contingency table analysis. (See Bishop et al., 1975:375–61; see also Press and Wilson, 1978).

The clear implication is that one should avoid continuous independent variables if the dependent variable is discrete; if they cannot be avoided, they must be collapsed; if there are only continuous independent variables, try discriminant analysis.

This is unnecessary. Logit models are estimable and interpretable when the explanatory variables are both continuous and categorical (Hanushek and Jackson, 1977:200–7). The citation of Press and Wilson (1978) is particularly puzzling, since that article demonstrates unambiguously that a variety of properties of the logistic regression model make it generally superior to the discriminant function model. Chief among these is the admission by logistic regression of nonnormality in the explanatory variables; this makes logit models especially attractive when one or more of the explanatory variables are dichotomous. Moreover, logistic regression turns out in their examples to be slightly better than discriminant analysis in the classification of data between the two states of the dependent variable. The same point is made by Fienberg (1980:109), who also concludes: "Efficient computer programs for fitting logistic response models have now removed cost as a barrier to an appropriate analysis of data not meeting the basic assumptions of the discriminant function analysis."

Swafford's (1980) confusion on this point seems to stem from an earlier section of his article (pp. 678–9), in which he states that although maximum likelihood (ML) estimates are superior to those obtained from generalized least squares with an estimated variance-covariance matrix, use of ML entails certain limiting trade-offs. In particular, he claims that estimation problems and interpretation difficulties restrict ML estimates to what Haberman (1978) and others term hierarchical log-linear models—models consisting only of discrete variables, where the restriction to zero of the parameters for a factor implies that all higher-order interactions involving that factor are necessarily zero as well. Since logit analysis with ordered or continuous explanatory variables does not correspond to a so-defined hierarchical linear model, Swafford says its use should be avoided.

The main problem here is Swafford's focus on hierarchical models. While we grant that models involving interaction terms but not main effects are difficult to interpret (Fienberg, 1978, 1980), it does not follow that models in-

* Direct all correspondence to Paul P. L. Cheung, Population Studies Center, University of Michigan, 1225 South University Avenue, Ann Arbor, Michigan 48109. We are grateful to William M. Mason for helpful comments.

volving continuous explanatory variables or mixtures of continuous and categorical variables are necessarily subject to this difficulty. The problem of interpretation to which Swafford alludes is really that of marginality (Nelder, 1974, 1976, 1977; Fienberg, 1978, 1980). Models possessing marginality are models in which the presence of higher-order terms dictates the inclusion of their constituent lower-order components.

This distinction becomes clearer with reference to Table 1, a scheme for classifying the models of discrete dependent variables according to marginality and type of explanatory variable. Cells B and D refer to models that are nonmarginal, clearly nonhierarchical, and thus fraught with the attendant interpretation difficulties mentioned by Swafford (1980:679). (See Evers and Namboodiri, 1978, for an example of a Cell B model.) Cell A is Swafford (1980) and Haberman's (1978) definition of a hierarchical-log linear model. But Swafford claims that Cell C should be relegated to discriminant analysis, or (in certain circumstances) to generalized least squares with scaled explanatory variables, or that it should be made into a Cell A problem by collapsing over the continuous explanatory variables. Since, however, these are still marginal models, interpretation of their parameters should be no problem, as a subsequent example will attest. With respect to the asserted difficulties in estimation (Swafford, 1980:679), it should simply be noted that workable, indeed, highly convenient solutions do exist.

Since logistic regression is possible given continuous independent variables, the only rationales for categorizing or collapsing them should involve either substantive criteria or better specification of the functional form of the equation. In arbitrarily collapsing variables for the sake of the supposed limitations of logistic regression, the true relationship between the underlying continuous variable and the dependent quantity may be distorted (Hanushek and Jackson, 1977:200).

Table 1. Typology of Models Involving Discrete Dependent Variables

Type of Independent Variable	Marginality: Marginal	Marginality: Non-Marginal
Discrete	A	B
Continuous or Mixed	C	D

A good example of this distortion may be found in Press and Wilson's (1978) data on state-level population change. Each of the 50 states is assigned a 0 if its percent change in population between 1960 and 1970 was below the median percent change for all states, and a 1 if the percent change was above the median. Explanatory variables include per capita income (in \$1000), the birth rate (per 100), the death rate (per 100), the absence or presence of a coastline (0 or 1), and degree of urbanization (0 if a state is less than 70% urban, 1 otherwise). Summary statistics for these variables are found in Table 2; the data are not reproduced here since they are readily available elsewhere (Press and Wilson, 1978:Table 3; Fienberg, 1980:Table 6–6).

First, the logistic regression model

$$\ln(\Omega) = \lambda_0 + \lambda^1 \text{ Income} + \lambda_2 \text{ Births} + \lambda_3 \text{ Deaths} + \lambda_4\text{Coast} + \lambda_5 \text{ Urban} \qquad (1)$$

is estimated, where Ω is the odds of a state experiencing population growth that is above the median state-growth rate, and the independent variables are as described above. The results are displayed in the first column of Table 3. High per capita income, high birth rates, and the presence of a coastline are all positively associated with population growth; death rates have a negative association, as does the level of urbanization, although this last effect is not statistically significant. These results are consistent with previous results based on the analysis of subsets of these data (Press and Wilson, 1978). The interpretation of the logit coefficients for the continuous independent variables is straightforward: An increase of one unit of per capita income (\$1000) results in an increase of 2.974 in the log-odds of above-the-median population growth; more felicitously, it multiplies the odds of population growth by a factor of 19.57—the same sort of description that is offered by Swafford (1980:672) for categorical independent variables.

But what if—following Swafford (1980:686)—the analyst felt "compelled . . . to categorize continuous variables for the sake of contingency table analysis"; would the results be comparable to those obtained using the underlying continuous independent variables? The answer depends on which categorization scheme is chosen. If the motivation for this transformation stems solely from a desire to alleviate the imagined estimation problem, then a simple dichotomization seems a likely choice. One seemingly reasonable, if arbitrary, classification involves dividing per capita income at \$3700, the birth rate at 1.9, and the death rate at 1.0. The resulting distributions of these variables are found in Table 2, where the suffix (D) denotes the dichotomous version.

The second column of Table 3 shows the coefficients of a second estimation of Equation 1, where this time all the independent variables

Table 2. Minima, Maxima, Means and Standard Deviations for Per Capita Income, Birth Rate, Death Rate, Presence of Coastline and Urbanization: 50 States

	Minimum	Maximum	Mean	Standard Deviation
Income (C)	2.626	4.917	3.737	.570
Income (D)	0	1	.58	.499
Births (C)	1.5	2.7	1.874	.244
Births (D)	0	1	.44	.501
Deaths (C)	0.5	2.4	.966	.247
Deaths (D)	0	1	.48	.505
Coast	0	1	.48	.505
Urban	0	1	.42	.499

(C) continuous
(D) dichotomized

are dichotomous.[1] The signs of the coefficients are almost the same as in Model I—only the insignificant effect of urbanization is reversed—but whereas previously four of the five variables had a significant impact on population change, now only two (coast and deaths) have coefficients approaching twice their standard errors. Both models fit the data comparably, but in the second case (Model II) the analyst might be lead to exclude terms for per capita income and birth rate, and to arrive at some substantially different conclusions.

The point here is not that Model I is superior to Model II, nor that continuous independent variables are always to be preferred over their categorical versions.[2] It is simply that they may differ, and since *both* are permissable within the framework of a logistic regression model, the choice between them should not be made on the mistaken grounds of "technique" limitations.

Table 3. Logistic Regression of Population Change on Per Capita Income, Birth Rate, Death Rate, Presence of Coastline and Urbanization: 50 States

Variables	Model I	Model II
Constant	−12.98 (8.231)	−0.4755 (1.100)
Income (C)	2.974[a] (1.368)	
Income (D)		1.347 (0.969)
Births (C)	4.809[a] (2.203)	
Births (D)		0.2111 (0.945)
Deaths (C)	−7.853[a] (3.485)	
Deaths (D)		−3.086[a] (0.942)
Coast	1.604[a] (0.801)	1.831[a] (0.874)
Urban	−1.000 (1.118)	0.3865 (0.832)
G^2	44.80	42.35
d.f.	44	44
Baseline G^2	69.31(d.f.=49)	

[a] Coefficient greater than twice its standard error.
(C) continuous
(D) discrete

[1] Since Swafford (1980) argues that collapsing must be done for the sake of contingency table analysis, the reader may wonder why Model II has 44 degrees of freedom (50 cases −6 parameters). The explanation is that this is a binomial response model comparable to a disaggregated contingency table. Alternatively, a 32-cell table could have been set up in which the binomial response in each cell is the summation of all cases with similar scores on the five independent variables. Since in fact ten of these cells would contain no cases, this table would be fitted with 16 d.f. (32 cells − 6 parameters − 10 zero cells) and have a G^2 of 21.93. The corresponding baseline model has a G^2 of 48.90; the difference between these two is equal to the difference between the G^2 of Model II and its baseline model (48.90 − 21.93 = 26.97; 69.31 − 42.35 = 26.96; both differences are distributed as G^2 with 5 d.f.). Still another alternative would be to fit a log-linear model to the further unfolded 64-cell table, but the results would still be the same. Parameter estimates for either version of the binomial response model are identical.

[2] In fact, if per capita income is trichotomized, it can be seen that low incomes significantly depress the propensity for population change, while there is no additional gain for being among the high income states. Similarly, Model I can be substantially improved by the inclusion of a second-order term for per capita income.

REFERENCES

Baker, R. J. and J. A. Nelder
1978 The GLIM System (Release 3). Oxford, U.K.: Numerical Algorithms Group.

Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul W. Holland
1975 Discrete Multivariate Analysis: Theory and Practice. Cambridge, Massachusetts: MIT Press.

Evers, Mark and N. Krishnan Namboodiri
1978 "On the design matrix strategy in the analysis of categorical data." Pp. 86–111 in Karl F. Schuessler, ed., Sociological Methodology 1979. San Francisco: Jossey-Bass.

Fienberg, Stephen E.
1978 "A note on fitting and interpreting parameters in models for categorical data." Pp. 112–8 in Karl F. Schuessler, ed., Sociological Methodology 1979. San Francisco: Jossey-Bass.
1980 The Analysis of Cross-Classified Data. 2nd edition. Cambridge, Massachusetts: MIT Press.
Haberman, Shelby J.
1978 Analysis of Qualitative Data. Volume 1. New York: Academic Press.
Hanushek, Eric A. and John E. Jackson
1977 Statistical Methods for Social Scientists. New York: Academic Press.
Nelder, J. A.
1974 "Analysis of variance programs, least squares, and square tables (letter to the editor)." Applied Statistics 23:232.
1976 "Hypothesis testing in linear models (letter to the editor)." American Statistician 30:101.
1977 "A reformulation of linear models (with discussion)." Journal of the Royal Statistical Society, series A 140:48–77.
Nelder, J. A. and R. W. M. Wedderburn
1972 "Generalized linear models." Journal of the Royal Statistical Society, series A 135:370–84.
Press, S. James and Sandra Wilson
1978 "Choosing between logistic regression and discriminant analysis." Journal of the American Statistical Association 73:699–705.
Swafford, Michael
1980 "Three parametric techniques for contingency table analysis: a nontechnical commentary." American Sociological Review 45:664–90.

LOGISTIC REGRESSION*

(REPLY TO CHEUNG AND SMITH)

MICHAEL SWAFFORD
Vanderbilt University

Although Cheung and Smith have misread my position on hierarchical modeling, they are quite justified in pointing to the availability of logistic regression.[1] Since contingency tables (of frequencies) by definition involve only categoric variables, logistic regression seemed beyond the scope of my already lengthy article when I wrote it. Logistic regression also seemed too problematic to warrant inclusion in a nontechnical article. In one of the most authoritative treatises on log-linear modeling, for example, Bishop et al. (1975:358–9; 371–2) had described some aspects of logistic regression as only "occasionally of a manageable form." And Fienberg did not see fit to make his claim (quoted by Cheung and Smith) about the availability of efficient computer programs until the 1980 edition of his book. We are fortunate that so much progress on logistic regression has been made recently, and that an appropriate computer program has been introduced into a statistical package generally available to sociologists.[2] I would agree that, given a mixture of continuous and categoric explanatory variables, one now ought to consider logistic regression in situations where the dependent variable can legitimately be dichotomized. Furthermore, even when all explanatory variables are continuous, logistic regression may constitute a useful alternative to discriminant analysis if the assumptions underlying the latter are violated and the dependent variable is dichotomous.

Regrettably, instead of merely pointing out the existence of this new statistical tool and illustrating its use, Cheung and Smith devote most of their comment to speculating erroneously on why I did not mention it and attacking a distorted representation of my views on hierarchical modeling. They begin their third paragraph by summarizing my views as follows: ". . . although maximum likelihood (ML) estimates are superior to those obtained from generalized least squares . . . use of ML entails certain limiting trade-offs." In the first place, though ML estimates may well be superior, I did not make that claim. In fact, I said (p. 679) that "one should not take for granted the superiority of ML estimates. . . ." More important, I did not claim that the use of ML estimates necessarily entails trade-offs. The gist of page 679 was that, if one relied on a computer program like Goodman's ECTA to obtain ML estimates, the use of such estimates would entail trade-offs because ECTA limits analysis to hierarchical models. This seemed to

* Address all communications to Michael Swafford, Department of Sociology and Anthropology, Vanderbilt University, Nashville, Tennessee 37235.

[1] Some readers may be perplexed by the fact that "logistic" and "logit" seem to be used interchangeably by Cheung and Smith. In this response, I am using the terms as Bishop et al. (1975:357–61) and Fienberg (1980:97–105) use them. "Logistic" and "logit" both refer to models in which the criterion variable is the log of the odds of falling into one category of the dependent variable rather than the other. However, "logistic" is usually employed in reference to models in which the explanatory variables may be continuous. The special case of a logistic model in which the explanatory variables are categoric is equivalent to a logit model.

[2] See BMDPLR in Dixon et al. (1981). Incidentally, the same series contains BMDP4F, a very attractive program for log-linear modeling that was omitted from fn. 25 of my article. Also, a widely distributed program for linear probability models which should have been included in fn. 25 is FUNCAT, in the SAS statistical package available to IBM users.

be an important point, both because ECTA is by far the most commonly used program for log-linear modeling among sociologists, and because even referees of an earlier version of the article summarily dismissed the GSK technique as inferior to ECTA on the sole grounds that it did not produce ML estimates.

Cheung and Smith continue explicating my views as follows: ". . . he claims that estimation problems and interpretation difficulties restrict ML estimates to what Haberman (1978) and others term hierarchical log-linear models. . . ." Nowhere did I state or imply this viewpoint. All I said (p. 679) was this: "The emphasis on hierarchical models is usually explained by the fact that ML estimates for nonhierarchical models are more difficult to calculate than those for hierarchical models (Haberman, 1978 [161]) and because they are often difficult to interpret substantively (Bishop et al., 1975 [38])." This is a far cry from claiming that ML estimates are restricted to hierarchical models. Furthermore, my claim about difficulties in interpretation did not even pretend to concern logistic regression, a subject I did not broach in the article. Thus, Cheung and Smith's criticism of my "asserted" or "imagined" difficulties takes my statement out of context. My statement, incidentally, remains true: The reason that hierarchical log-linear models were given so much attention was that the required calculations and interpretation of coefficients were thereby simplified.

Finally, the third paragraph misrepresents my reasoning as follows: "Since logit analysis with ordered or continuous explanatory variables does not correspond to a so-defined hierarchical linear model, Swafford says its use should be avoided." Again, nowhere did I state or imply that nonhierarchical models (logit or otherwise) should be avoided. The bulk of page 679, in fact, illustrates the need for them in certain cases, and footnote 25 lists at least two computer programs (by Magidson and by Haberman) with which nonhierarchical log-linear or logit modeling could be performed.

If I had space to render a sentence-by-sentence reply, I could point out several other similar unjustified misinterpretations. Suffice it to say, it is clear from the term "seems" in the third paragraph that Cheung and Smith were speculating as to why I did not mention logistic regression. This erroneous speculation then serves as the basis for much of the comment. However, my reason for omitting logistic regression was not that I had mistakenly set up the hierarchy principle as the standard by which to judge models; nor was I "confused." I simply considered logistic regression too problematic and inaccessible at the time I wrote the article.

As for the example provided by Cheung and Smith, those who are unfamiliar with logistic regression will no doubt find it heuristically useful. Nevertheless, three notes of caution seem warranted. First, the coefficients in their Table 3 would have been more accurately described as "direct effects" or "partial effects" than as "associations"; likewise, a proper interpretation of their coefficients would have included "controlling on other variables in the equation." Second, without special justification there is no reason for performing tests of significance when the entire population (that is, 50 states) is being analyzed. Third, their enthusiasm for logistic regression should be tempered somewhat by the fact that, as Fienberg himself has pointed out (1980:104), some problems inherent in logistic regression still remain unsolved. Still, as Cheung and Smith have correctly pointed out, logistic regression constitutes a promising new tool in our statistical workshops.

REFERENCES

Bishop, Yvonne M. M., Stephen E. Fienberg and Paul W. Holland
1975 Discrete Multivariate Analysis: Theory and Practice. Cambridge, Massachusetts: MIT Press.

Dixon, W. J., M. B. Brown, L. Engelman, J. W. Frane, M. A. Hill, R. I. Jennrich and J. D. Toporek (eds.)
1981 BMDP Statistical Software 1981. Los Angeles: University of California Press.

Fienberg, Stephen E.
1980 The Analysis of Cross-Classified Categorical Data. Second edition. Cambridge, Massachusetts: MIT Press.

Haberman, Shelby J.
1978 Analysis of Qualitative Data. Vol. 1. New York: Academic Press.

REVIEWERS FOR THE 1981 ISSUE

Abrahamson, Mark
Adams, Bert M.
Akers, Ronald L.
Albrecht, Gary
Albrecht, Stan L.
Alexander, Jeffrey
Alexander, Karl
Alwin, Duane F.
Arney, William
Allen, Michael P.
Althauser, Robert P.
Anderson, James
Anderson, Ronald
Arce, Carlos
Armer, J. Michael
Barnes, Douglas
Baron, James
Bean, Frank D.
Beck, E. M.
Beegle, J. Allan
Bell, Wendell
Booth, Alan
Benson, Kenneth
Berk, Richard A.
Berg, Ivar
Bianchi, Suzanne
Bibb, Robert
Bielby, William
Black, Donald
Blau, Peter M.
Bogdan, Robert
Bohrnstedt, George
Bonacich, Edna
Booth, Alan
Bordua, David
Bredo, Eric
Breed, Warren
Brent, Edward
Bumpass, Larry
Bunker, Stephen G.
Burawoy, Michael
Burt, Ronald S.
Cahalan, Donald
Cain, Pamela S.
Camilleri, Santo Frank
Campbell, Richard T.
Cannon, Lynn
Carey, James
Champion, Dean J.
Chase, Ivan
Chase-Dunn, Christopher
Cherlin, Andrew
Chilton, Roland J.
Choldin, Harvey
Clark, John P.
Clark, Terry N.
Coatsworth, John
Cockerham, William
Cohen, Albert
Cohen, Lawrence.
Cole, Stephen
Collins, Randall
Collver, O. Andrew
Comaroff, John
Cook, Karen
Cook, Martha
Corsaro, William
Coverman, Shelley
Crain, Robert
Cramer, James C.
Cutright, Phillips
Danigelis, Nicholas L.
Darrow, William
Davidson, James D.
Deegan, Mary J.
Demerath, N. J. III
Denzin, Norman K.
Dinitz, Simon
Dohrenwend, Bruce
Dornbusch, Sanford M.
Driedger, Leo
Dutton, Diana
Dworkin, Roger
Elder, Glen Holl
Emerson, Richard M.
Ennis, Phillip
Erickson, Eugene C.
Erlanger, Howard S.
Evans, Peter
Fahey, Anthony
Farley, Reynolds
Faunce, William A.
Feagin, Joe R.
Featherman, David L.
Felson, Marcus
Fendrich, James M.
Ferber, Marianne
Ferree, Myra Marx
Fine, Gary A.
Firebaugh, Glenn
Fishman, Joshua
Fliegel, Frederick
Fligstein, Neil
Form, William
Freeman, Howard
Freeman, John
Frisbie, Parker
Fuguitt, Glenn V.
Gagnon, John
Gamson, William
Garnier, Maurice
Geis, Gilbert L.
Gelles, Richard
Geschwender, James
Gibbs, Jack P.
Giddens, Anthony
Glenn, Norval D.
Glynn, Evelyn
Goffman, Erving
Gordon, Scott
Gorecki, Jan
Gottfredson, Gary D.
Gottfredson, Michael R.
Gove, Walter R.
Granovetter, Mark
Greeley, Andrew M.
Greenberg, David
Griffin, Larry J.
Grimshaw, Allen D.
Groves, Robert M.
Grusky, Oscar
Guest, Avery M.
Hage, Jerald
Hagen, John
Hagstrom, Warren O.
Hall, Peter M.
Hamblin, Robert L.
Hardesty, Monica
Hauser, Robert M.
Hawkins, Richard
Hawley, Amos H.
Hazelrigg, Lawrence
Heer, David
Henshel, Richard
Hicks, Alex
Higgins, Paul
Hirsch, Paul
Hirschi, Travis W.
Hirschman, Charles
Hogan, Dennis P.
Holmstrom, Linda
Hood, Thomas C.
Horan, Patrick M.
Hout, Michael
Huber, Joan
Hughes, Michael
Imrey, Peter B.
Inkeles, Alex
Inverarity, James
Irwin, John
Isaac, Larry
Jackman, Mary R.
Jackman, Robert W.
Jackson, Elton
Jacobs, David
Jencks, Christopher S.
Johnson, Barclay
Jones, Robert A.
Jurik, Nancy
Kalleberg, Arne L.
Karsh, Bernard
Kaufman, Robert
Kelly, William
Kerckhoff, Alan C.
Kessler, Ronald
Kilgore, Sally
Kimberly, James C.
Kleck, Gary
Knoke, David
Kohn, Melvin
Kohout, Frank
Komorita, Samuel
Krohn, Marvin
Kuhn, Alfred
Ladinsky, Jack
LaFree, Gary
Land, Kenneth C.
Lantz, Herman
Laslett, Barbara
Lauderdale, Patrick
Lawler, Edward
Leik, Robert
Lempert, Richard O.
Levin, Donald
Lewis, J. David
Lieberson, Stanley
Liebert, Roland
Light, Donald
Lin, Nan
Lincoln, James R.
Lingoes, James C.

Liska, Allen
Lizotte, Alan
Loftin, Colin K.
Logan, John R.
Long, Larry
Lueschen, Guenther
Lundman, Richard
Maddox, George L.
Maines, David R.
Marcus, Philip
Mare, Robert D.
Maris, Ronald
Marsden, Peter
Marsh, Robert
Marshall, Harvey
Marshall, James
Marx, Gary T.
Mason, Karen O.
Massey, Douglas S.
Mayhew, Leon
Mazur, Allan
McPhail, Clark
McCarthy, John
McCleary, Richard
McKee, McClendon
McLaughlin, Steven D.
McMahon, Anne M.
McMillen, David
McNall, Scott G.
McNamee, Steven
Mechanic, David
Mehan, Hugh
Meier, John
Meier, Robert F.
Meyer, John W.
Meyer, Marshall W.
Micklin, Michael
Miller, Joanne
Milner, Murray
Mirowsky, John
Modigliani, Andre
Moen, Phyllis
Moore, Gwen
Mueller, Charles W.
Mullins, Nicholas C.
Myers, Martha
Nagel, Ilene H.
Nagel, Stuart S.
Namboodiri, N. K.
Newby, Howard
Nielsen, Francois
Nolan, Patrick
Ornstein, Michael
Palloni, Alberto
Palmore, Erdman
Pampel, Fred
Parcel, Toby L.
Parker, Robert
Payne, David
Pearlin, Leonard
Peterson, Richard
Pettigrew, Thomas F.
Phillips, David
Ploch, Donald R.
Pope, Whitney
Popenoe, David
Poston, Dudley L.
Powers, Charles
Pratt, Lois
Pu-lum, Thomas
Quadnagno, Jill
Quinn, Robert
Ragin, Charles
Ratcliff, Richard E.
Rau, William
Rehberg, Richard
Reiss, Albert J.
Reiss, Ira L.
Riley, Matilda W.
Robin, Ellen
Robinson, Jerry W.
Robinson, Robert
Roman, Paul
Roncek, Dennis
Roof, W. Clark
Rosenbaum, James
Rosenfeld, Rachel
Ross, Catherine E.
Rossi, Alice S.
Rossi, Peter H.
Roth, Guenther
Rothschild-Whitt, Joyce
Roy, William G.
Rubin, Lillian
Rudwick, Elliott
Ryder, Norman B.
Salancik, Gerald
Scanzoni, John
Scarr, Sandra
Schervish, Paul
Schoen, Robert
Schoenberg, Ronald
Schooler, Carmi
Schuessler, Karl F.
Schuman, Howard
Schur, Edwin
Schwartz, Mildred A.
Schwirian, Kent P.
Scotch, Norman A.
Seeman, Melvin
Segal, David R.
Selznick, Philip
Shephard, Jon
Short, James F.
Shover, Neal
Simon, Rita J.
Simon, William
Skocpol, Theda
Sly, David
Smith, Doug
Smith, Joel
Smith-Lovin, Lynn
Snow, David
Somers, Robert H.
Southwood, Kenneth
Spaeth, Joe
Spenner, Kenneth
Spitze, Glenna D.
Steffensmeier, Darrell
Stern, Robert
Stinchcombe, Arthur
Stolzenberg, Ross M.
Stryker, Sheldon
Summers, Gene
Suter, Larry E.
Swafford, Michael
Sykes, Richard
Szelenyi, Ivan
Tallman, Irving
Thielens, Wagner
Thomlinson, Ralph
Tibbetts, Paul
Tickamayer, Ann
Tilly, Charles H.
Tilly, Louise
Treas, Judith
Turk, Austin T.
Tyree, Andrea
Useem, Michael
Useem, John
Vanneman, Reeve
Villemez, Wayne J.
Vogel, Ezra
Wahrman, Ralph
Waite, Linda J.
Walton, John
Wallerstein, Immanuel
Warriner, Charles
Wax, Murray L.
Webster, Murray
Weigart, Andrew
Weinberg, Martin S.
Weinstein, Deena
Wheaton, Blair
Whyte, Martin K.
Wiebe, Paul
Williams, Robin M.
Wilsnack, Richard W.
Wilson, James Q.
Wilson, Kenneth
Winship, Christophuer
Wisan, Gail
Wohlstein, Ronald
Wolfgang, Marvin E.
Wood, James R.
Wright, James
Wuthnow, Robert
Wyner, Gordon
Yinger, J. Milton
Young, Frank W.
Zaret, David
Zelditch, Morris
Zerubavel, Eviatar
Zurcher, Louis A.

INDEX

AMERICAN SOCIOLOGICAL REVIEW

Published Bimonthly by the American Sociological Association

CONTENTS OF VOLUME 46, NUMBERS 1–6

ARTICLES

RESEARCH NOTES

COMMENTS, REPLIES, AND CORRECTIONS

U.S. POSTAL SERVICE

STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION

(Required by 39 U.S.C. 3685)

1. TITLE OF PUBLICATION	A. PUBLICATION NO.	2. DATE OF FILING
AMERICAN SOCIOLOGICAL REVIEW	0 0 0 3 1 2 2 4	9/30/81

3. FREQUENCY OF ISSUE	A. NO. OF ISSUES PUBLISHED ANNUALLY	B. ANNUAL SUBSCRIPTION PRICE
Bimonthly	6	$18

4. COMPLETE MAILING ADDRESS OF KNOWN OFFICE OF PUBLICATION (Street, City, County, State and ZIP Code) (Not printers)

1722 N Street, N.W., Washington, D.C. 20036

5. COMPLETE MAILING ADDRESS OF THE HEADQUARTERS OR GENERAL BUSINESS OFFICES OF THE PUBLISHERS (Not printer)

1722 N Street, N.W., Washington, D.C. 20036

6. FULL NAMES AND COMPLETE MAILING ADDRESS OF PUBLISHER, EDITOR, AND MANAGING EDITOR (This item MUST NOT be blank)

PUBLISHER (Name and Complete Mailing Address)

American Sociological Association, 1722 N Street, N.W., Washington, D.C. 20036

EDITOR (Name and Complete Mailing Address)

William H. Form, Department of Sociology, University of Illinois, Urbana, IL 61801

MANAGING EDITOR (Name and Complete Mailing Address)

7. OWNER (If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given. If the publication is published by a nonprofit organization, its name and address must be stated.) (Item must be completed)

FULL NAME	COMPLETE MAILING ADDRESS
American Sociological Association	1722 N Street, N.W., Washington, D.C. 20036

8. KNOWN BONDHOLDERS, MORTGAGEES, AND OTHER SECURITY HOLDERS OWNING OR HOLDING 1 PERCENT OR MORE OF TOTAL AMOUNT OF BONDS, MORTGAGES OR OTHER SECURITIES (If there are none, so state)

FULL NAME	COMPLETE MAILING ADDRESS
None	

9. FOR COMPLETION BY NONPROFIT ORGANIZATIONS AUTHORIZED TO MAIL AT SPECIAL RATES (Section 411.3, DMM only)
The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes (Check one)

- [x] (1) HAS NOT CHANGED DURING PRECEDING 12 MONTHS
- [ ] (2) HAS CHANGED DURING PRECEDING 12 MONTHS

(If changed, publisher must submit explanation of change with this statement.)

10. EXTENT AND NATURE OF CIRCULATION	AVERAGE NO. COPIES EACH ISSUE DURING PRECEDING 12 MONTHS	ACTUAL NO. COPIES OF SINGLE ISSUE PUBLISHED NEAREST TO FILING DATE
A. TOTAL NO. COPIES (Net Press Run)	16,000	16,000
B. PAID CIRCULATION 1. SALES THROUGH DEALERS AND CARRIERS, STREET VENDORS AND COUNTER SALES	NONE	NONE
2. MAIL SUBSCRIPTION	13,486	14,234
C. TOTAL PAID CIRCULATION (Sum of 10B1 and 10B2)	13,486	14,234
D. FREE DISTRIBUTION BY MAIL, CARRIER OR OTHER MEANS SAMPLES, COMPLIMENTARY, AND OTHER FREE COPIES	NONE	NONE
E. TOTAL DISTRIBUTION (Sum of C and D)	13,486	14,234
F. COPIES NOT DISTRIBUTED 1. OFFICE USE, LEFT OVER, UNACCOUNTED, SPOILED AFTER PRINTING	2,514	1,766
2. RETURN FROM NEWS AGENTS	NONE	NONE
G. TOTAL (Sum of E, F1 and 2 - should equal net press run shown in A)	16,000	16,000

11. I certify that the statements made by me above are correct and complete

SIGNATURE AND TITLE OF EDITOR, PUBLISHER, BUSINESS MANAGER, OR OWNER

Jann Ruckel Administrative Officer

PS Form 3526 (Page 1)
June 1980

(See instruction on reverse)

Notice to Contributors

A processing fee of $10 is required for each paper submitted, except for comments; such fees will be waived for student members of ASA. A check or money order, made payable to the American Sociological Association, should accompany each submission. The fee must be paid in order for the review process to begin.

The length of an *ASR* paper is generally ten (10) printed pages or thirty (30) typed pages. However, the *ASR* is happy to accept manuscripts up to fifty (50) typed pages in length for review. We also are interested in receiving manuscripts for a Research and Notes Section.

To permit anonymity in the review of manuscripts, keep identifying material out of the manuscript. Attach a cover page giving authorship, institutional affiliation and acknowledgments, and provide only the title as identification on the manuscript and abstract.

ASA Multiple Submissions Policy: Submission of a manuscript to a professional journal clearly implies commitment to publish in that journal. The competition for journal space requires a great deal of time and effort on the part of editorial readers whose only compensation for this service is the opportunity to read papers prior to publication and the gratification associated with discharge of professional obligation. For these reasons, *the American Sociological Association regards submission of a manuscript to a professional journal while that paper is under review by another journal as unacceptable.*

SUBMISSION AND PREPARATION OF MANUSCRIPTS

1. Submit four (4) copies and retain the original for your files. Copies may be Xerox, mimeograph or multilith, but not carbons.
2. Enclose a stamped self-addressed postcard for acknowledgment of manuscript receipt. *Manuscripts will not be returned.* Please do not send your original copy.
3. All copy must be typed, doublespaced (including indented material, footnotes and references) on 8½ by 11 inch white opaque paper. Lines must not exceed six (6) inches. Margins must be a minimum of one inch.
4. Include four (4) copies of an abstract of no more than 150 words.
5. Type each table on a separate page. Insert a location note at the appropriate place in the text, e.g., "Table 2 about here." Send copies, retain originals.
6. Figures must be drawn in India ink on white paper. Send copies, retain originals.
7. Clarify all symbols with notes in the margins of the manuscript. Circle these and all other explanatory notes not intended for printing.
8. Footnotes should not be placed in the body of the text. Type them (doublespaced) and attach them as a separate appendix to the text. Number them consecutively throughout the text. Footnotes are to be kept to a minimum and used only for substantive observations. Source citations are made within the text rather than in footnotes.
9. Acknowledgments, credits and grant numbers are placed on the title page with an asterisk.

REFERENCE FORMAT

A. *In the text:* All source references are to be identified at the appropriate point in the text by the last name of the author, year of publication and pagination where needed. Identify subsequent citations of the same source in the same way as the first, not using "*ibid.*," "*op. cit.*," *or* "*loc. cit.*" Examples follow:

1. If author's name is in the text, follow it with year in parentheses. [". . . Duncan (1959) . . ."]
2. If author's name is not in the text, insert, in parentheses, the last name and year, separated by a comma. [". . . (cf. Gouldner, 1963) . . ."]
3. Pagination (without "p." or "pp.") follows year of publication after a colon. [". . . Kuhn (1970:71)."]
4. Give both last names for dual authors; for more than two use "et al." in the text. When two authors have the same last name, use identifying initials in the text. For institutional authorship, supply minimum identification from the beginning of the complete citation. [". . . (U.S. Bureau of the Census, 1963:117) . . ."]
5. Separate a series of references with semicolons and enclose them within a single pair of parentheses. [". . . (Burgess, 1968; Marwell et al., 1971; Cohen, 1962) . . ."]

B. *In the appendix:* List all source citations by author, and within author by year of publication, in an appendix titled "References." The reference appendix must be complete and include all references in the text. The use of "et al." is not acceptable in the appendix; list the names of all authors. (See A. 4. for text format.)
If there is more than one reference to the same author and year, distinguish them by the letters, a, b, etc. added to the year. [". . . Levy (1965a:331) . . ."] Give the publisher's name in as brief a form as is fully intelligible. For example, John A. Wiley and Sons should be "Wiley." If the cited material is unpublished, use "forthcoming" with name of journal or publisher; otherwise use "unpublished."
Use no underlining, italics or abbreviations.
Examples follow:

1. *Books:*
   Jud, Gerald J., Edgar W. Mills, Jr. and Genevieve Walters Burch
   1970 Ex-Pastors. Philadelphia: Pilgrim Press.
   U.S. Bureau of the Census
   1960 Characteristics of Population. Volume I. Washington, D.C.: U.S. Government Printing Office.
   Bernard, Claude
   [1865] 1957 An Introduction to the Study of Experimental Medicine. Tr. Henry Copley Greene, New York: Dover.
2. *Periodicals:*
   Conger, Rand
   Forthcoming "The effects of positive feedback on direction and amount of verbalization in a social setting." Pacific Sociological Review.
   Merton, Robert K.
   1963a "The ambivalence of scientists." Bulletin of The Johns Hopkins Hospital 112:77–97.
   1963b "Resistance to the systematic study of multiple discoveries in science." European Journal of Sociology 4:237–82.
3. *Collections:*
   Davie, M.
   1938 "The pattern of urban growth." Pp. 133–61 in G. Murdock (ed.), Studies in the Science of Society. New Haven: Yale University Press.

See recent issues for further examples.

Revised 1980 ASR